

Exakte kleinste Quadrate Interpolierende
Konvergenzresultate vom Korovkin-Typ und Anwendungen im
Kontext der numerischen Approximation von Erhaltungsgleichungen

Von der Carl-Friedrich-Gauß-Fakultät
der Technischen Universität Carolo-Wilhelmina zu Braunschweig

zur Erlangung des Grades

Doktor der Naturwissenschaften (Dr. rer. nat.)

genehmigte

Dissertation

von Dipl.-Math. Oliver Nowak

geboren am 30. Juli 1980

in Helmstedt

Eingereicht am	6. Mai 2009
Mündliche Prüfung am	29. Juni 2009
Referent	Professor Dr. Thomas Sonar
Korreferent	Professor Dr. Klaus-Jürgen Förster



(2009)

Für meine Eltern in tiefer Dankbarkeit.

Zusammenfassung.

In den frühen 1980er Jahren veröffentlichten Peter Lancaster und Kestutis Šalkauskas die so genannte *moving least squares interpolation*, eine Methode zur glatten Interpolation von beliebig verteilten multivariaten Daten. Die entsprechenden interpolierenden Funktionen minimieren lokal singular gewichtete kleinste Quadrate Funktionale. Diese Arbeit untersucht die *exakte kleinste Quadrate Interpolation*, ein zur Methode nach Lancaster und Šalkauskas verwandter Ansatz. Nach einem Vergleich dieser beiden Ansätze beweisen wir Konvergenzresultate für die exakte kleinste Quadrate Interpolation vom Korovkin-Typ. Hierfür übertragen und verallgemeinern wir bekannte Konvergenzsätze vom Korovkin-Typ für lineare und positive Operatoren auf Folgen von linearen und regulären Operatoren, die auf Räumen stetiger Funktionen definiert sind. Neben diesen Untersuchungen bringen wir Anwendungen der exakte kleinste Quadrate Interpolation in einem Kontext von Ableitungs-Approximationen durch Finite Differenzen und der numerischen Behandlung von nichtlinearen skalaren Erhaltungsgleichungen.

Abstract.

In the early 1980s, Peter Lancaster and Kestutis Šalkauskas introduced *moving least squares interpolation*. This is a technique for the smooth interpolation of multivariate scattered data. The corresponding interpolating functions are local minimizers of certain least squares functionals equipped with singular weighting functions. In this thesis we examine *exact least squares interpolation*, which is a technique related to the approach given by Lancaster and Šalkauskas. We provide a comparison of the two approaches. Moreover, we prove Korovkin-type convergence results for exact least squares interpolation. For this reason, we establish a generalised Korovkin-type theorem for sequences of linear and regular operators on spaces of continuous functions utilizing common Korovkin-type theorems for linear and positive operators. We also study basic applications from exact least squares interpolation in the context of derivative approximations by finite difference operators and the numerical treatment of nonlinear scalar conservation laws.

Danksagungen.

Die vorliegende Arbeit entstand während meiner Tätigkeit als wissenschaftlicher Mitarbeiter und Assistent in der Arbeitsgruppe Partielle Differentialgleichungen des Instituts Computational Mathematics an der TU Braunschweig.

Mein erster Dank gebührt von ganzem Herzen meinem Mentor Professor Dr. Thomas Sonar für die große Chance, mir diese Arbeit ermöglicht und über die Zeit fachlich und insbesondere menschlich begleitet zu haben.

Herrn Professor Dr. Klaus-Jürgen Förster danke ich für die Übernahme des Korreferats.

Allen Mitgliedern und insbesondere Frau Ines Engling als «Chefin» der Arbeitsgruppe Partielle Differentialgleichungen gilt mein Dank für die Zusammenarbeit und eine äußerst angenehme und für mich sicherlich unvergessliche Assistenten-Zeit an der Carolo-Wilhelmina.

Darüber hinaus möchte ich meinen Kollegen Frau Dipl.-Math. Martina Pawlak und Herrn Dipl.-Math. Dipl.-Phys. Thomas Hennecke für die Durchsicht einer frühen Version dieser Arbeit danken.

Meinem ehemaligen Kollegen Herrn Dr. Tim Scharlau danke ich nicht nur für die Durchsicht von Teilen dieser Arbeit, sondern auch für die Abende mit Diskussion und Zigarre während und nach unserer gemeinsamen Assistenten-Zeit.

Meinen Eltern und meinem Bruder danke ich für alle Unterstützung hinweg über die vielen Jahre.

Diese Arbeit ist meinen Eltern Frau Ingeborg Nowak und Herrn Udo Nowak in tiefer Dankbarkeit gewidmet.

Inhaltsverzeichnis

Einleitung	Seite 1
 Kapitel 1	 Seite 5
Kleinste Quadrate Interpolation	
1.1 Über univariate Polynom-Interpolation und kleinste Quadrate Approximation	5
1.2 Bestapproximation in Räumen mit innerem Produkt	8
1.3 Multivariate Interpolation nach Shepard	9
1.4 Verallgemeinerte kleinste Quadrate Interpolation	13
1.5 Ein Vergleich von $LS-\mathcal{G}f$ und $\mathcal{G}f$	22
1.6 Eine Kardinalbasis-Darstellung der exakte kleinste Quadrate Interpolierenden	26
 Kapitel 2	 Seite 29
Exakte kleinste Quadrate Interpolierende im Kontext von Ableitungs-Approximationen	
2.1 Zur Berechnung von Ableitungen von exakte kleinste Quadrate Interpolierenden	29
2.2 Exakte kleinste Quadrate Interpolierende zur Konstruktion von Finite Differenzen ...	34
2.3 Inverse Euklidische Normen als positive Interpolationsgewichte	38
2.4 Die Konstruktion und Analyse symmetrischer 3-Punkte Finite Differenzen	39

Kapitel 3 **Seite 45**
Konvergenzresultate für exakte kleinste Quadrate Interpolierende vom Korovkin-Typ

3.1 Konvergenz der univariaten Shepard-Interpolation	46
3.2 Über die Namensgebung des Konvergenzsatzes von Bohman und Korovkin	53
3.3 Konvergenz von Folgen fast positiver linearer Operatoren auf $C(K; \mathbb{R})$	54
3.4 Das Konvergenzproblem für die exakte kleinste Quadrate Interpolation	59
3.5 Konvergenz bei konstanter monomialer Interpolationsbasis	61
3.6 Eine detailliertere Untersuchung der Kernfunktionen	64
3.7 Konvergenz bei quadratischer monomialer Interpolationsbasis	69
3.8 Konvergenz bei linearer monomialer Interpolationsbasis	73

Kapitel 4 **Seite 79**
Ein Resultat von simultaner Approximation

4.1 Simultane Approximation von konvexen Operatoren	79
4.2 Exkurs in die Theorie der Sobolev-Räume	81
4.3 Simultane Approximation durch kompakte Einbettungen	82

Kapitel 5 **Seite 85**
Anwendungen im Kontext der numerischen Approximation von Erhaltungsgleichungen

5.1 Konservative Verfahren für die nichtlineare Advektionsgleichung	85
5.2 Ein konservatives Verfahren aus exakte kleinste Quadrate Interpolierenden	87
5.3 Ein informeller Exkurs über Upwind-Verfahren	94
5.4 Ad hoc Upwind-Verfahren durch Variation der Parameter des Interpolationsgewichtes	98
5.5 Numerische Beispiele	99

Literaturverzeichnis **Seite 107**

‘What we know is a drop.
What we don’t know is an ocean.’
Sir Isaac Newton (1643-1727)

Einleitung

Diese Arbeit handelt von der *exakten kleinsten Quadrate Interpolation*, einer Methode zur Interpolation von beliebig verteilten multivariaten Daten. Wie der Untertitel «Konvergenzresultate vom Korovkin-Typ und Anwendungen im Kontext der numerischen Approximation von Erhaltungsgleichungen» vermuten lässt, betrachten wir die exakte kleinste Quadrate Interpolation aus zwei verschiedenen Blickwinkeln. Zunächst widmen wir uns den mathematischen Grundlagen und Konvergenzaspekten des reinen Interpolations-Ansatzes.

In den späten 1960er Jahren veröffentlichte Donald Shepard in [26] eine Methode zur Interpolation von Datenpunkten im \mathbb{R}^3 . Shepards Ansatz basiert auf einer Darstellung vom Lagrange-Typ für die entsprechende Interpolierende. Für jeden der Datenpunkte $(x_i, y_i, f_i) \in \mathbb{R}^{2+1}$ ist demnach eine Funktion φ_i mit der Eigenschaft $\varphi_i(x_i, y_i) = 1$ und $\varphi_i(x_j, y_j) = 0$ für beliebige andere Datenpunkte $(x_j, y_j, f_j) \neq (x_i, y_i, f_i)$ zu konstruieren. Shepard nutzte

$$\varphi_i(x, y) = \frac{(x - x_i)^2 + (y - y_i)^2}{\sum_{j \neq i} ((x - x_j)^2 + (y - y_j)^2)},$$

so dass die gesuchte Interpolierende durch

$$\mathbb{R}^2 \ni (x, y) \mapsto \sum_i f_i \cdot \frac{(x - x_i)^2 + (y - y_i)^2}{\sum_{j \neq i} ((x - x_j)^2 + (y - y_j)^2)}.$$

gegeben ist. Ein wesentlicher Vorteil dieses Ansatzes ist die einfach durchzuführende multivariate Verallgemeinerung: Für gegebene Datenpunkte $(x_j, f_j) \in \mathbb{R}^d \times \mathbb{R}$ ist das multivariate Analogon zur Shepard-Interpolierenden

$$\mathbb{R}^d \ni x \mapsto \sum_i f_i \cdot \frac{\|x - x_i\|_2^2}{\sum_{j \neq i} \|x - x_j\|_2^2},$$

wobei $\|\cdot\|_2$ die Euklidische Norm auf \mathbb{R}^d bezeichnet. Die Einfachheit dieser Darstellung der Shepard-Interpolierenden bringt allerdings auch einen wesentlichen Nachteil mit sich, das so genannte *flat spot-Phänomen*. Hiermit wird das Verschwinden des Gradienten der Interpolierenden an allen Punkten (x_i, y_i) bezeichnet, das auf der speziellen Wahl der φ_i basiert.

Innerhalb des ersten Kapitels «*Kleinste Quadrate Interpolation*» bringen wir zunächst eine Verallgemeinerung der Shepard-Interpolation, die so genannte *moving least squares interpolation*. Diese wurde zu Beginn der 1980er Jahre von Peter Lancaster und Kestutis Šalkauskas in [17] vorgeschlagen. Deren Ansatz unterliegt im Allgemeinen nicht dem flat spot-Phänomen und

beinhaltet die Shepard-Interpolation als Spezialfall. Zu einer vorgelegten Menge von Datenpunkten (x_i, f_i) ist die kleinste Quadrate Interpolierende $x \mapsto \mathcal{L}_x f(x)$ nach Lancaster und Šalkauskas eine Funktion, die lokal in x ein singulär gewichtetes kleinste Quadrate Funktional von der Form

$$\mathcal{L}_x f \mapsto \sum_j (\mathcal{L}_x f(x_j) - f_j)^2 \frac{1}{\|x - x_j\|_2^\alpha}$$

minimiert. Rigoros betrachtet ist dieses Funktional, infolge der singulären Gewichte $\|x - x_j\|_2^{-\alpha}$, zunächst für jedes $x = x_j$ undefiniert. Betrachten wir davon unbeeindruckt folgende Heuristik: Angenommen, $\mathcal{L}_x f(x)$ minimiert das kleinste Quadrate Funktional für beliebiges x . Das heißt, der Wert

$$\sum_j (\mathcal{L}_x f(x_j) - f_j)^2 \frac{1}{\|x - x_j\|_2^\alpha}$$

existiert insbesondere für $x = x_j$. Dann folgt aufgrund von $\lim_{x \rightarrow x_j} \|x - x_j\|_2^{-\alpha} = \infty$ schon

$$\lim_{x \rightarrow x_j} (\mathcal{L}_x f(x_j) - f_j)^2 = 0.$$

Dies impliziert die Interpolationseigenschaft der Interpolierenden $x \mapsto \mathcal{L}_x f(x)$, das heißt

$$\mathcal{L}_{x_j} f(x_j) = f_j.$$

Unabhängig von dieser heuristischen Betrachtungsweise stellt sich die Frage, ob und in welchem Sinn die Minimierung des kleinste Quadrate Funktional auch für $x = x_j$ mathematisch rigoros zu handhaben ist. Diese Minimierung ist äquivalent zur Lösung eines korrespondierenden Systems von Normalengleichungen, das in Analogie zum kleinste Quadrate Funktional für den Fall $x = x_j$ zunächst nicht definiert ist. Die auftretenden Gleichungen enthalten „ ∞ “-Einträge. Lancaster und Šalkauskas schlagen eine Restriktion des Lösungsraums dieser Normalengleichungen vor, auf deren Basis diese „ ∞ “-Einträge unter Einbeziehung der Theorie der multivariaten Shepard-Interpolation verschwinden.

Ein anderer Ansatz zur Handhabung der singulären Normalengleichungen wurde in [28] von Thomas Sonar und Co-Autoren als Verallgemeinerung einer Idee von Kunle [16] gegeben. Sonar betrachtet regularisierte Versionen der singulären Gewichte, was eine Regularisierung der Normalengleichungen impliziert und deren Lösbarkeit garantiert. Ferner stellt sich heraus, dass die Lösungen auch im Fall der „verschwindenden Regularisierung“ existieren. Diesen Ansatz bezeichnen wir als *exakte kleinste Quadrate Interpolierende*.

Einem Vergleich der beiden Ansätze anschließend leiten wir eine *Kardinalbasis-Darstellung* von exakte kleinste Quadrate Interpolierenden in Analogie zur Darstellung von Shepard-Interpolierenden her. Diese nutzen wir insbesondere innerhalb des dritten Kapitels im Kontext des Konvergenzproblems für die exakte kleinste Quadrate Interpolation.

Es stellt sich heraus, dass exakte kleinste Quadrate Interpolierende beliebig glatt konstruiert werden können. Basierend auf dieser Tatsache, beschäftigen wir uns innerhalb des zweiten Kapitels *«Exakte kleinste Quadrate Interpolierende im Kontext von Ableitungs-Approximationen»* mit der Berechnung von Ableitungswerten von exakte kleinste Quadrate Interpolierenden und geben einen ersten Anwendungsbezug im Kontext der Approximation von Ableitungswerten von differenzierbaren Funktionen durch Finite Differenzen.

Angenommen, es existiert eine stetige Funktion f mit $f(x_i) = f_i$ für die Menge der zu interpolierenden Datenpunkte (x_i, f_i) . Dann stellt sich die Frage, ob die aus diesen Daten bestimmte exakte kleinste Quadrate Interpolierende in einem speziellen Sinn gegen die Funktion f konvergiert, falls die Anzahl der Datenpunkte kontinuierlich erhöht wird. Das dritte Kapitel über *«Konvergenzresultate für exakte kleinste Quadrate Interpolierende vom Korovkin-Typ»* ist dieser Fragestellung gewidmet. Wir beantworten diese Frage mit Hilfe der zur exakte kleinste Quadrate Interpolation gehörenden Approximations-Operatoren, die Räume von stetigen Funktionen auf Kompakta in sich selbst abbilden. Innerhalb dieser Betrachtungsweise kann die Shepard-Interpolation als positiver Operator aufgefasst werden, das heißt positive Funktionen werden auf positive Funktionen abgebildet.

Für Folgen von positiven und linearen Operatoren existieren *Konvergenzresultate vom Korovkin-Typ* [3], die wir direkt auf das Konvergenzproblem für die Shepard-Interpolation anwenden. Die durch die exakte kleinste Quadrate Interpolation gegebenen Approximations-Operatoren sind im Allgemeinen nicht positiv. Jedoch stellen sich diese Operatoren als *regulär* heraus, das heißt sie sind als Differenz von zwei positiven Operatoren darstellbar. Für spezielle Folgen von regulären und linearen Operatoren, so genannte *Folgen von fast positiven linearen Operatoren*, verallgemeinern wir klassische Konvergenzresultate vom Korovkin-Typ. Dies ermöglicht es uns, Konvergenzaussagen für die exakte kleinste Quadrate Interpolation zu beweisen.


Im vierten Kapitel *«Ein Resultat von simultaner Approximation»* gehen wir für den univariaten Fall der Frage nach, ob Resultate ähnlich denen in Kapitel 3 auch für Ableitungen von exakte kleinste Quadrate Interpolierenden Gültigkeit besitzen. Wir beweisen diese Frage betreffend ein spezielles Ergebnis, das auf kompakten Einbettungen von Sobolev-Räumen basiert.

Mit dem fünften und abschließenden Kapitel dieser Arbeit über *«Anwendungen im Kontext der numerischen Approximation von Erhaltungsgleichungen»* knüpfen wir erneut an die Resultate über die Konstruktion von Finite Differenzen in Kapitel 2 an. Hierbei konstruieren wir Verfahren zur numerischen Approximation von linearen und nichtlinearen Advektionsgleichungen und legen ein besonderes Augenmerk auf die Konstruktion von Upwind-Verfahren.

Kapitel 1

Kleinste Quadrate Interpolation

1.1	Über univariate Polynom-Interpolation und kleinste Quadrate Approximation	5
1.2	Bestapproximation in Räumen mit innerem Produkt	8
1.3	Multivariate Interpolation nach Shepard	9
1.4	Verallgemeinerte kleinste Quadrate Interpolation	13
1.5	Ein Vergleich von $LS-\mathcal{G}f$ und $\mathcal{G}f$	22
1.6	Eine Kardinalbasis-Darstellung der exakte kleinste Quadrate Interpolierenden	26

ieses Kapitel ist der Interpolation von beliebig verteilten multivariaten Daten mit Hilfe so genannter *kleinste Quadrate Interpolierender* gewidmet. *Beliebig verteilt* bezeichnet hierbei die Tatsache, dass die zugrundeliegenden Daten nicht notwendigerweise auf einem entsprechend regulären Gitter liegen. Wir führen diese Interpolations-Methoden dem historischen Kontext entsprechend als natürliche Erweiterung eines Interpolationsansatzes nach Shepard [26] aus dem Jahre 1968 ein und bringen die grundlegenden Definitionen der kleinste Quadrate Interpolation wie Lancaster und Šalkauskas in der Arbeit [17]. Neben der gerade erwähnten klassischen Methode nach Lancaster und Šalkauskas bringen wir eine neuere und leicht modifizierte Version dieses Ansatzes, die auf Ideen von Sonar et al. [28] basiert und die wir als *exakte kleinste Quadrate Interpolation* bezeichnen. Das Kapitel schließt mit einem detaillierten Vergleich der beiden Ansätze. Wir beginnen dieses Kapitel mit einem Exkurs in die Theorie der univariaten Polynom-Interpolation, sowie einer Zusammenfassung der Theorie der Bestapproximation in Räumen mit innerem Produkt.

1.1 Über univariate Polynom-Interpolation und kleinste Quadrate Approximation

Wir bringen zunächst einige klassische Resultate für univariate Polynom-Interpolation sowie kleinste Quadrate Methoden, die auf Polynom-Approximation basieren.

1.1.i Univariate Polynom-Interpolation

Eine Problemstellung: Seien $(x_1, y_1), \dots, (x_n, y_n) \in \mathbb{R} \times \mathbb{R}$ mit $x_i \neq x_j$ für $i \neq j$. Konstruiere eine Polynomfunktion $p : \mathbb{R} \rightarrow \mathbb{R}$ mit der Eigenschaft

$$y_1 = p(x_1), \dots, y_n = p(x_n). \quad (1.1)$$

Wir wollen hier vereinfachend den Begriff des Polynoms gleichwertig mit dem der Polynomfunktion benutzen. Ein Polynom mit der Eigenschaft (1.1) wird *Interpolationspolynom* für die Datenmenge

$$\{(x_1, y_1), \dots, (x_n, y_n)\}$$

genannt.

Satz 1.1 Seien $(x_1, y_1), \dots, (x_n, y_n) \in \mathbb{R} \times \mathbb{R}$ mit $x_i \neq x_j$ für $i \neq j$. Dann existiert ein eindeutig bestimmtes Polynom von Grade $n - 1$ mit

$$y_i = p(x_i), \quad 1 \leq i \leq n,$$

und es gilt

$$p(x) = \sum_{i=1}^n y_i \cdot \prod_{k \neq i} \frac{x - x_k}{x_i - x_k} \quad (1.2)$$

für alle $x \in \mathbb{R}$.

Der Beweis dieses Standardresultates ist in nahezu jedem Buch über Numerische Analysis zu finden. Verwiesen sei an dieser Stelle zum Beispiel auf [11].

Die Darstellung (1.2) wird *Lagrange-Form* des Interpolationspolynoms genannt.

Ist dem Interpolationsprozess ein Datenpaar $(x_{n+1}, y_{n+1}) \in \mathbb{R} \times \mathbb{R}$ hinzuzufügen wird deutlich, dass die so genannte Lagrange-Interpolation nicht Methode der Wahl ist, denn das gesamte Polynom (1.2) ist neu zu berechnen. Diese Prozedur lässt sich vermeiden, indem für die Darstellung des Interpolationspolynoms die *Newton-Form* (1.3) gewählt wird.

Satz 1.2 Das eindeutig bestimmte Interpolationspolynom in Satz 1.1 lässt sich ebenfalls durch

$$p(x) = \sum_{i=1}^n \left([[x_1, \dots, x_i]] \cdot \prod_{k=1}^{i-1} (x - x_k) \right), \quad (1.3)$$

mit

$$[[x_i]] = y_i, \quad (1.4a)$$

$$[[x_i, \dots, x_{i+k}]] = \frac{[[x_{i+1}, \dots, x_{i+k}]] - [[x_i, \dots, x_{i+k-1}]]}{x_{i+k} - x_i}. \quad (1.4b)$$

darstellen.

Die Ausdrücke in den Gleichungen (1.4a) und (1.4b) werden *dividierte Differenzen* genannt. Existiert eine Funktion mit $y_i = f(x_i)$, schreiben wir für die dividierten Differenzen $[[x_i; f]]$ bzw. $[[x_i, \dots, x_{i+k}; f]]$.

Bleiben wir zunächst im univariaten Fall, so ist oftmals die Approximation von gewissen Daten (an dieser Stelle wollen wir nicht näher auf den Sinn, in welchem diese zu verstehen ist, eingehen) einer Interpolation dieser durch Polynome vorzuziehen. Letztere neigen in bestimmten Fällen zu starken Oszillationen bei wachsender Datenanzahl. Hierbei spricht man vom so genannten *Runge-Phänomen* [24].

Wir wollen uns an dieser Stelle mit einem kurzen historischen Exkurs auf den Begriff der *kleinste Quadrate Approximation* einstellen. Im Jahre 1801 gelang es Carl Friedrich Gauß den Orbit des Planetoiden Ceres mit einer derart hohen Genauigkeit zu berechnen, dass es aufgrund dieser möglich war, Ceres an vorhergesagter Himmelsstelle erneut aufzufinden. Zu dieser Zeit war den Astronomen über einen großen Zeitraum nach Ceres' Entdeckung durch Piazzi zum Jahreswechsel 1800/01 eine Wiederentdeckung verwehrt geblieben. Innerhalb seiner Berechnungen nutzte Gauß seinerzeit eine *kleinste Quadrate Methode* zur Minimierung von Messfehlern [8].

1.1.ii Kleinste Quadrate Methoden basierend auf Polynom-Approximation

Eine Problemstellung: Seien $(x_1, y_1), \dots, (x_n, y_n) \in \mathbb{R} \times \mathbb{R}$ mit $x_i \neq x_j$ für $i \neq j$ und $U \subset C[a, b]$ ein Unterraum. Konstruiere eine Funktion $g \in U$ mit der Eigenschaft

$$\sum_{i=1}^n (y_i - g(x_i))^2 \leq \sum_{i=1}^n (y_i - f(x_i))^2 \quad \forall f \in U. \quad (1.5)$$

Verwenden wir an dieser Stelle den Unterraum $U = \text{span}\{e_0, e_1\}$, mit $e_i(x) = x^i$ führt dies zu den bekannten linearen Regressionsgeraden: Die minimierende Funktion g in (1.5) berechnet sich mit Hilfe von

$$\mathcal{R}(a, b) = \sum_{i=1}^n (y_i - (ax_i + b))^2 \longrightarrow \min!, \quad (1.6)$$

wobei wir g als $g(x) = \tilde{a}x + \tilde{b}$ mit $\mathcal{R}(\tilde{a}, \tilde{b}) = \min_{a, b \in \mathbb{R}} \mathcal{R}(a, b)$ erhalten. Mit Standardtechniken lassen sich die bekannten Werte

$$\tilde{a} = \frac{n \sum_{i=1}^n x_i y_i - \left(\sum_{i=1}^n x_i \right) \left(\sum_{i=1}^n y_i \right)}{n \sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i \right)^2}, \quad \tilde{b} = \frac{\left(\sum_{i=1}^n y_i \right) \left(\sum_{i=1}^n x_i^2 \right) - \left(\sum_{i=1}^n x_i \right) \left(\sum_{i=1}^n x_i y_i \right)}{n \sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i \right)^2} \quad (1.7)$$

ermitteln. Hierbei wird dann $y = g(x) = \tilde{a}x + \tilde{b}$ als *lineare Regressionsgerade* bezeichnet.

Die Approximation mit $U = \text{span}\{e_0, e_1, e_2\}$ ist äquivalent zum Minimierungsproblem

$$\mathcal{R}(a, b, c) = \sum_{i=1}^n (y_i - (ax_i^2 + bx_i + c))^2 \longrightarrow \min!.$$

Untersuchen wir Gleichung (1.6) noch einmal näher. Der \mathbb{R}^n sei hierfür mit dem gewöhnlichen Euklidischen Skalarprodukt (\cdot, \cdot) ausgestattet, wodurch die Euklidische Norm $\|\cdot\|_2$ auf natürliche Weise induziert wird. Ferner definieren wir Vektoren $p_{a,b}, y \in \mathbb{R}^n$ vermöge

$$p_{a,b} = (ax_1 + b, \dots, ax_n + b), \quad y = (y_1, \dots, y_n).$$

Dann schreibt sich (1.6) als

$$\|y - p_{a,b}\|_2^2 \longrightarrow \min!, \quad (1.8)$$

das heißt wir suchen insbesondere die *Bestapproximation* des Vektors y im Unterraum

$$\mathcal{P}_1 = \text{span}\{p_{a,b} \mid a, b \in \mathbb{R}\} \subset \mathbb{R}^n. \quad (1.9)$$

Werfen wir einen Blick auf die hinter diesem Ansatz stehende Approximationstheorie.

1.2 Bestapproximation in Räumen mit innerem Produkt

Dieser Abschnitt stellt die wesentlichen Ergebnisse über Bestapproximation in Räumen mit innerem Produkt dar. Da es sich um allgemein bekannte Standardresultate handelt, verzichten wir auf die Angabe von Beweisen und verweisen auf die Monographie [11].

Zunächst geben wir eine Definition von Bestapproximationen in normierten Räumen in Analogie zu Gleichung (1.5).

Definition 1.1 Sei $(V, \|\cdot\|)$ ein normierter Vektorraum und $U \subset V$ ein Unterraum. Ein Vektor $\tilde{u} \in U$ wird *Bestapproximation von $v \in V$ in U* genannt, falls die Ungleichung

$$\|\tilde{u} - v\| \leq \|u - v\|$$

für alle $u \in U$ gilt.

Theorem 1.1 Sei $(V, (\cdot, \cdot))$ ein \mathbb{K} -Vektorraum mit innerem Produkt, $U \subset V$ ein Unterraum von endlicher Dimension und $v \in V$. Dann existiert eine eindeutig bestimmte Bestapproximation $\tilde{u} \in U$ von v in U und es gilt

$$(\tilde{u} - v, u) = 0 \quad \forall u \in U. \quad (1.10)$$

Vice versa ist jedes $\tilde{u} \in U$ mit (1.10) eine eindeutig bestimmte Bestapproximation von v in U .

Der zweite Teil von Theorem 1.1 liefert einen Ansatz um Bestapproximationen via der Lösung von linearen Gleichungssystemen zu bestimmen.

Satz 1.3 Sei $U = \text{span}\{u_1, \dots, u_n\}$, $v \in V$ und $\tilde{u} = \sum_{i=1}^n \alpha_i u_i$ für Skalare $\alpha_1, \dots, \alpha_n \in \mathbb{K}$.

Dann ist \tilde{u} eine Bestapproximation von v in U , genau dann wenn

$$\sum_{i=1}^n \alpha_i (u_i, u_j) = (v, u_j) \quad (1.11)$$

für $1 \leq j \leq n$ gilt.

Beweis. Anwendung von (1.10) auf die Bestapproximation \tilde{u} sowie $u_1, \dots, u_n \in U$ liefert

$$0 = \left(\sum_{i=1}^n \alpha_i u_i - v, u_j \right) = \sum_{i=1}^n \alpha_i (u_i, u_j) - (v, u_j), \quad 1 \leq j \leq n.$$

Vice versa schließen wir (1.10) aus $(\tilde{u} - v, u_j) = 0$ für $1 \leq j \leq n$, denn $\{u_1, \dots, u_n\}$ ist eine Basis von U nach Definition. \square

Die Gleichungen (1.11) werden *Normalgleichungen* genannt. Die Lösung des Systems (1.11) ist eindeutig, solange u_1, \dots, u_n linear unabhängig sind, da in diesem Fall die *Gramsche Matrix*

$$((u_i, u_j))_{i,j=1}^n$$

regulär ist [11].

Wenden wir uns noch einmal dem Problem (1.6) bzw. (1.8) zu. Dieses kann mit Blick auf die Notation in Definition 1.1 auch wie folgt formuliert werden: *Gesucht ist die Bestapproximation $\tilde{u} \in \mathbb{R}^n$ von $(y_1, \dots, y_n) \in \mathbb{R}^n$ im Unterraum*

$$U = \text{span} \left\{ \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} \right\}.$$

Nach Satz 1.3 ist dann

$$\tilde{u} = \alpha \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} + \beta \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix},$$

mit Skalaren α, β , die durch das Gleichungssystem

$$\alpha \sum_{i=1}^n x_i^2 + \beta \sum_{i=1}^n x_i = \sum_{i=1}^n x_i y_i, \quad (1.12a)$$

$$\alpha \sum_{i=1}^n x_i + \beta \sum_{i=1}^n 1 = \sum_{i=1}^n y_i, \quad (1.12b)$$

gegeben sind. Die Lösung der Gleichungen (1.12a) und (1.12b) liefert erneut (1.7).

Wir können zeigen, dass die Bestapproximation im letzten Beispiel für den nicht-ausgearteten Fall (d.h. $\nexists a, b : y_i = ax_i + b \forall i$) keine Interpolation der gesamten Daten $(x_1, y_1), \dots, (x_n, y_n)$ mit sich bringt. Denn eine einfache Rechnung von Hand zeigt, dass $\tilde{\alpha}, \tilde{\beta}$ in (1.7)

$$\frac{1}{n} \sum_{i=1}^n y_i = \tilde{\alpha} \cdot \frac{1}{n} \sum_{i=1}^n x_i + \tilde{\beta}$$

erfüllen.

Die folgenden Abschnitte beschäftigen sich mit der Konstruktion einer Approximation, die auch mit einer Interpolationseigenschaft für eine gewisse Datenmenge ausgestattet ist.

1.3 Multivariate Interpolation nach Shepard

Satz 1.1 liefert sowohl die Existenz als auch die Eindeutigkeit der Lösung der univariaten Polynom-Interpolationsaufgabe. Basierend auf einer Idee von Shepard [26] für den zweidimensionalen Fall, bringen wir in diesem Abschnitt eine Methode zur Interpolation von beliebig verteilten multivariaten Daten

$$(x_1, f_1), \dots, (x_n, f_n) \in K \times \mathbb{R}, \quad (1.13)$$

für ein einfach zusammenhängendes Kompaktum $K \subset \mathbb{R}^d$ und $x_i \neq x_j$ für $i \neq j$. Diese Methode ist auch in leicht modifizierter Darstellung bei Lancaster und Šalkauskas [17] zu finden. Nehmen wir hierfür die Darstellung

$$\sum_{i=1}^n f_i \varphi_i \quad (1.14)$$

der Interpolierenden an, so sind die Funktionen φ_i mit einer Kronecker- δ -Eigenschaft

$$\varphi_i(x_j) = \delta_{ij}$$

auszustatten, um die Interpolationseigenschaft für (1.13) zu gewährleisten.

Sei nun $\alpha > 1$ eine natürliche Zahl und bezeichne $w_\alpha : \mathbb{R}^d \rightarrow \mathbb{R}_{>0} \cup \{\infty\}$ die im Ursprung singuläre Funktion

$$w_\alpha(x) = \|x\|_2^{-\alpha}. \quad (1.15)$$

Für $x \in \mathbb{R}^d$ definieren wir (insbesondere durch Grenzwertbildung für $x = x_j$) die Funktionen $\varphi_1, \dots, \varphi_n : \mathbb{R}^d \rightarrow \mathbb{R}$ durch

$$\varphi_i(x) = \frac{w_\alpha(x - x_i)}{\sum_{j=1}^n w_\alpha(x - x_j)}, \quad 1 \leq i \leq n. \quad (1.16)$$

Die in (1.16) definierten Funktionen φ_i , welche wir im weiteren Verlauf als *Kernfunktionen* bezeichnen wollen, haben die folgenden Eigenschaften:

Satz 1.4 Die Funktionen $\varphi_1, \dots, \varphi_n$ gemäß (1.16) sind beliebig glatt, das heißt $\varphi_i \in C^\infty(\mathbb{R}^d; \mathbb{R})$, und erfüllen

- (a) $\varphi_i(x_j) = \delta_{ij}$ für alle $i, j = 1, \dots, n$,
- (b) $0 \leq \varphi_i(x) \leq 1$ für alle $x \in \mathbb{R}^d$,
- (c) $\varphi_i(x) = 0$ genau dann wenn $x = x_j$ mit $i \neq j$,
- (d) $\sum_{i=1}^n \varphi_i(x) = 1$ für alle $x \in \mathbb{R}^d$,
- (e) $D^\mu \varphi_i(x)|_{x=x_j} = 0$ für beliebige Multi-Indizes μ von der Ordnung $|\mu| = 1$ und $1 \leq i, j \leq n$.

Beweis. Den Beweis von $\varphi_i \in C^\infty(\mathbb{R}^d; \mathbb{R})$ stellen wir zunächst zurück und verweisen zunächst auf Lemma 2.1. Die Punkte (b) und (d) sind trivialerweise erfüllt. Die Identität

$$\varphi_i(x) = \frac{w_\alpha(x - x_i)}{\sum_{j=1}^n w_\alpha(x - x_j)} = \frac{1}{1 + \sum_{\substack{j=1 \\ j \neq i}}^n \frac{w_\alpha(x - x_j)}{w_\alpha(x - x_i)}}$$

liefert (a) sowie (c) unter Beachtung der Singularität von $w_\alpha(\cdot - x_j)$ bei $x = x_j$. Das so genannte *flat spot*-Phänomen (e) ist unter Einbezug der Differenzierbarkeit der φ_i eine Konsequenz von (a) und (b). \square

Mit Hilfe der Eigenschaften der φ_i nach Satz 1.4 kann die Konstruktion einer Interpolierenden mit Darstellung (1.14) völlig problemlos durchgeführt werden.

Definition 1.2 Für festes $\alpha > 1$ und beliebig verteilte Interpolationsdaten

$$(x_1, f_1), \dots, (x_n, f_n) \in K \times \mathbb{R}, \quad K \subset \mathbb{R}^d, \quad (1.17)$$

heißt die Funktion $\mathcal{S}_n^\alpha[x_{1 \rightarrow n} | f_{1 \rightarrow n}] : \mathbb{R}^n \rightarrow C^\infty(K; \mathbb{R})$, definiert durch

$$\mathcal{S}_n^\alpha[x_{1 \rightarrow n} | f_{1 \rightarrow n}](x) = \sum_{i=1}^n f_i \varphi_i(x), \quad x \in K,$$

mit den φ_i gemäß (1.16), die *Shepard-Interpolierende für die Daten* (1.17). Existiert ferner eine Funktion $f \in C(K; \mathbb{R})$ mit $f(x_i) = f_i$ für $1 \leq i \leq n$, schreiben wir für die Shepard-Interpolierende kurz $\mathcal{S}_n^\alpha f : C(K; \mathbb{R}) \rightarrow C^\infty(K; \mathbb{R})$ mit

$$\mathcal{S}_n^\alpha f(x) = \sum_{i=1}^n f(x_i) \varphi_i(x), \quad x \in K. \quad (1.18)$$

Hierbei beruht die Glattheit der Shepard-Interpolierenden in Definition 1.2 offensichtlich auf der Glattheit der Kernfunktionen (1.16).

Eine weitere Konsequenz der Darstellung (1.18) und Lemma 1.4 ist die *Stabilität* der Shepard-Interpolation im Sinne von

$$\min_{1 \leq i \leq n} f(x_i) \leq \mathcal{S}_n^\alpha f(x) \leq \max_{1 \leq i \leq n} f(x_i), \quad x \in K.$$

Die Auswirkungen des *flat spot*-Phänomens, also das Verschwinden der ersten Ableitungen der Interpolierenden bei sämtlichen x_i (siehe Punkt (e) in Satz 1.4), werden in [27] näher untersucht.

Wir bringen an dieser Stelle eine für die Notationen im Folgenden wesentliche Definition.

Definition 1.3 Sei $\varepsilon > 0$ und $g : \Omega \subseteq \mathbb{R}^d \rightarrow \mathbb{R}$ eine auf einer Teilmenge Ω von \mathbb{R}^d definierte Funktion. Dann setzen wir für $x \in \Omega$ den Funktionswert $g(x + \varepsilon) := g(x + \varepsilon \mathbb{1})$ für den Vektor $\mathbb{1} = (1, \dots, 1)^T \in \mathbb{R}^d$.

Einen tieferen Einblick in die Struktur der Shepard-Interpolation gibt die folgende Herleitung: Für festes $x \in K$ und $\varepsilon > 0$ sei $Q_{x,\varepsilon} : \mathbb{R} \rightarrow \mathbb{R}$ die punktweise definierte quadratische Funktion

$$Q_{x,\varepsilon}(a) = \sum_{j=1}^n (a - f(x_j))^2 w_\alpha(x - x_j + \varepsilon), \quad (1.19)$$

wobei durch das Addieren des $\varepsilon > 0$ im Argument der Funktion w_α in (1.19) der Wert $Q_{x,\varepsilon}(a)$ für beliebige $x \in K$ existiert. Ferner ist $Q_{x,\varepsilon}$ als Polynomfunktion auf ganz K stetig differenzierbar mit

$$\frac{d}{da} Q_{x,\varepsilon}(a) = 2 \sum_{j=1}^n (a - f(x_j)) w_\alpha(x - x_j + \varepsilon), \quad x \in K.$$

Das übliche notwendige Kriterium zur Bestimmung eines lokalen Minimums von $Q_{x,\varepsilon}$, das heißt $\frac{d}{da} Q_{x,\varepsilon}(a) = 0$, liefert

$$a \sum_{j=1}^n w_\alpha(x - x_j + \varepsilon) = \sum_{i=1}^n f(x_i) w_\alpha(x - x_i + \varepsilon) \iff a = \sum_{i=1}^n f(x_i) \frac{w_\alpha(x - x_i + \varepsilon)}{\sum_{j=1}^n w_\alpha(x - x_j + \varepsilon)}.$$

Mit $\frac{d^2}{da^2} Q_{x,\varepsilon}(a) = 2 \sum_{j=1}^n w_\alpha(x - x_j + \varepsilon) > 0$ folgt, dass es sich bei a um ein lokales, aufgrund der quadratischen Struktur von $Q_{x,\varepsilon}$ sogar um ein globales Minimum handelt. Für festes $x \in K$ konvergiert demnach das $Q_{x,\varepsilon}$ minimierende $a = a(x, \varepsilon)$ mit $\varepsilon \rightarrow 0$ gegen den Wert $\mathcal{S}_n^\alpha[x_{1 \rightarrow n} | f_{1 \rightarrow n}](x)$ bzw. $\mathcal{S}_n^\alpha f(x)$.

Diese Überlegungen legen es nahe die Shepard-Interpolierende zunächst ohne weitere Begriffsdefinition als *kleinste Quadrate Interpolierende* zu bezeichnen.

Wir beenden diesen Abschnitt mit einer weiteren Herleitung der Shepard-Interpolation. Für festes $x \in K$ und $\varepsilon > 0$ definieren wir für $g, h \in C(K; \mathbb{R})$

$$(g, h)_{x,\varepsilon} = \sum_{i=1}^n g(x_i) h(x_i) w_\alpha(x - x_i + \varepsilon). \quad (1.20)$$

Lemma 1.1 *Durch (1.20) ist für festes $x \in K$ und $\varepsilon > 0$ ein inneres Produkt auf dem Raum $C(K; \mathbb{R})$ gegeben.*

Beweis. Aufgrund von

$$(g, h)_{x,\varepsilon} = \sum_{i,j=1}^n g(x_i) h(x_j) \delta_{ij} w_\alpha(x - x_i + \varepsilon)$$

ist durch (1.20) eine positiv definite und symmetrische Bilinearform auf $C(K; \mathbb{R})$ gegeben. \square

Statten wir also $C(K; \mathbb{R})$ mit dem inneren Produkt $(\cdot, \cdot)_{x,\varepsilon}$ aus. Nach Theorem 1.1 ist die eindeutig bestimmte Bestapproximation $\tilde{f}_\varepsilon(x) e_0$ von $f \in C(K; \mathbb{R})$ bezüglich $(\cdot, \cdot)_{x,\varepsilon}$ im Unterraum $\text{span}\{e_0\}$ mit $e_0(x) = 1$ durch

$$(\tilde{f}_\varepsilon(x) e_0, e_0)_{x,\varepsilon} = (f, e_0)_{x,\varepsilon} \quad (1.21)$$

charakterisiert. Die Gleichung (1.21) nach $\tilde{f}_\varepsilon(x)$ aufgelöst liefert

$$\tilde{f}_\varepsilon(x) = \left(\sum_{j=1}^n w_\alpha(x - x_j + \varepsilon) \right)^{-1} \sum_{i=1}^n f(x_i) w_\alpha(x - x_i + \varepsilon),$$

und wir erhalten

$$\mathcal{S}_n^\alpha f(x) = \lim_{\varepsilon \rightarrow 0} \tilde{f}_\varepsilon(x), \quad x \in K,$$

als weitere Darstellung der Shepard-Interpolierenden.

Fassen wir die wesentlichen Ergebnisse dieses Abschnitts noch einmal zusammen: Die Shepard-Interpolation bietet die Möglichkeit der Interpolation von beliebig verteilten multivariaten Daten und lässt sich aus der Lösung eines quadratischen Variationsproblems bzw. als spezielle Bestapproximation der zu interpolierenden Daten gewinnen. Die Shepard-Interpolation unterliegt dem *flat spot*-Phänomen. In einem Kontext der simultanen Interpolation von Funktionswerten und Approximation bzw. Interpolation von Ableitungswerten, durch *simultane Approximation* bezeichnet, kann die Shepard-Interpolation demnach nicht verwendet werden. Im folgenden Abschnitt sehen wir, wie die oben aufgezeigte Bestapproximationsformulierung der Shepard-Interpolation die Möglichkeit zur Verallgemeinerung dieses Interpolationsansatzes liefert, die auch Anwendung im Kontext der simultanen Approximation finden wird.

1.4 Verallgemeinerte kleinste Quadrate Interpolation

Bisher haben wir den Begriff *kleinste Quadrate Interpolation* genutzt, ohne näher zu erläutern was rigoros unter einer solchen Methode zu verstehen ist. Zu Beginn dieses Abschnitts werden wir basierend auf den Formulierungen der Shepard-Interpolation eine genaue Definition liefern. Weiter sei angemerkt, dass von nun an für jede zu interpolierende Datenmenge

$$(x_1, f_1), \dots, (x_n, f_n) \in K \times \mathbb{R},$$

die Existenz einer Funktion $f \in C(K; \mathbb{R})$ mit $f(x_i) = f_i$ für $1 \leq i \leq n$ unter Berücksichtigung der Shepard-Interpolation angenommen wird.

1.4.i Verallgemeinerung der Interpolation nach Shepard

Werfen wir noch einmal einen Blick auf die Bestapproximationsformulierung der Shepard-Interpolation in Abschnitt 1.3. Dort stellte sich heraus, dass die Shepard-Interpolierende zu einer Funktion $f \in C(K; \mathbb{R})$ in $x \in K$ im Wesentlichen durch die Bestapproximation von f im Unterraum $\text{span}\{e_0\}$ bezüglich des inneren Produktes $(\cdot, \cdot)_x$ gemäß (1.20) gegeben ist. Eine natürliche Erweiterung dieser Methode ist offensichtlich durch Variation des betrachteten Unterraums gegeben. Zur näheren Spezifizierung dieser Erweiterung geben wir vorab einige Definitionen.

Definition 1.4 Eine Funktion $w \in C^l(\mathbb{R}^d \setminus \{0\}; \mathbb{R}_{>0})$ mit den Eigenschaften

$$\lim_{x \rightarrow 0} w(x) = +\infty,$$

$$\lim_{x \rightarrow x_k} \left(w(x - x_k) \frac{w(x - x_l)}{\sum_{i=1}^n w(x - x_i)} \right) \in \mathbb{R}, \quad 1 \leq k \neq l \leq n,$$

und

$$\lim_{\varepsilon \rightarrow 0} \frac{D^\mu w(\varepsilon)}{w(\varepsilon)^2} = 0$$

für jeden Multi-Index $\mu \in \mathbb{N}^d$ von der Ordnung $|\mu| = 1$ heißt *positives Interpolationsgewicht*.

Die Funktionen w_α in (1.15) sind positive Interpolationsgewichte nach Definition 1.4.

Definition 1.5 Eine Menge von Funktionen $\{b^{(1)}, \dots, b^{(k)}\}$ mit den Eigenschaften

- (i) $b^{(1)} \equiv 1$,
- (ii) $b^{(i)} \in C^m(K; \mathbb{R})$, $1 \leq i \leq k$,
- (iii) $\{b^{(i)}\}_{i=1}^k$ sind linear unabhängig über jeder Teilmenge von k Vektoren in $\{x_1, \dots, x_n\}$,

wird *Interpolationsbasis* genannt. Die Bedingung (iii) bedeutet hierbei, dass die Vektoren

$$\underline{b}(x_i) = (b^{(1)}(x_i), \dots, b^{(k)}(x_i))^T \in \mathbb{R}^k$$

linear unabhängig für k paarweise verschiedene Indizes i sind.

Wir nutzen in diesem Fall die Schreibweise $\underline{b}(x_i)$ um auszudrücken, dass der Vektor von den Funktionen $b^{(1)}, \dots, b^{(k)}$ abhängt. Bei der Shepard-Interpolation haben wir es mit dem Fall $k = 1$ mit $b^{(1)} = e_0$ und insbesondere $m = \infty$ zu tun. Dies ist aus gewissen Glattheitsaspekten für die im Folgenden einzuführende kleinste Quadrate Interpolierende ein optimaler Wert. Natürlich ist jedes $m \in \mathbb{N}$ ebenso zulässig.

1.4.ii Interpolierende kleinste Quadrate: Konstruktion nach Lancaster und Šalkauskas

Dieser Abschnitt führt die kleinste Quadrate Interpolation nach Lancaster und Šalkauskas [17] ein. Hierbei erweitern wir die Resultate der Originalarbeit auf allgemeine positive Interpolationsgewichte. Betrachten wir dazu noch einmal das innere Produkt (1.20) für den Fall $\varepsilon = 0$, das heißt

$$(g, h)_{x,0} \equiv (g, h)_x = \sum_{i=1}^n g(x_i)h(x_i)w_\alpha(x - x_i), \quad g, h \in C(K; \mathbb{R}). \quad (1.22)$$

Wählen wir hierbei $x = x_j \in \{x_1, \dots, x_n\}$, bildet $(\cdot, \cdot)_x$ im Allgemeinen kein inneres Produkt auf $C(K; \mathbb{R})$, denn einer der Summanden in (1.22) ist für $x = x_j$ singulär. Blenden wir dieses Problem einen Moment lang aus. Sei nun $f \in C(K; \mathbb{R})$ und $\{b^{(1)}, \dots, b^{(k)}\}$ eine Interpolationsbasis aus der wir für festes $x \in K$ einen Unterraum $\text{span}\{\tilde{b}_x^{(1)}, \dots, \tilde{b}_x^{(k)}\} \subset C(K; \mathbb{R})$ vermöge der konstanten Funktion

$$\tilde{b}_x^{(1)} = \left(\sum_{j=1}^n w(x - x_j) \right)^{-\frac{1}{2}}, \quad (1.23)$$

sowie den Funktionen

$$\tilde{b}_x^{(i)} = b^{(i)} - \mathcal{S}_n[w]b^{(i)}(x), \quad 2 \leq i \leq k, \quad (1.24)$$

konstruieren. Hierbei bezeichnet $\mathcal{S}_n[w]$ den auf positive Interpolationsgewichte erweiterten Shepard-Interpolationsoperator, das heißt

$$\mathcal{S}_n[w]f(x) = \sum_{i=1}^n f(x_i)\phi_i(x), \quad x \in K. \quad (1.25)$$

Für die Funktionen $\phi_i : K \rightarrow \mathbb{R}$ in (1.25), die punktweise für $x \in K$ durch

$$\phi_i(x) = \frac{w(x - x_i)}{\sum_{j=1}^n w(x - x_j)} \quad (1.26)$$

definiert sind, kann ein zu Satz 1.4 (und den darin beschriebenen Funktionen φ_i) nahezu analoges Resultat bewiesen werden. Einzig und allein unterscheiden sich die Differentiationsklassen der ϕ_i und φ_i : Aufgrund von $w \in C^l(\mathbb{R}^d \setminus \{0\}; \mathbb{R}_{>0})$ gilt $\phi_i \in C^l(\mathbb{R}^d; \mathbb{R})$.

Definition 1.6 Sei $f \in C(K; \mathbb{R})$ beliebig gewählt, $\{b^{(1)}, \dots, b^{(k)}\}$ eine Interpolationsbasis und $w \in C^l(\mathbb{R}^d \setminus \{0\}; \mathbb{R}_{>0})$ ein positives Interpolationsgewicht. Für festes $x \in K$ bezeichne

$$\text{LS-}\mathcal{L}_x f : K \rightarrow \mathbb{R}$$

die Bestapproximation von f in $\text{span}\{\tilde{b}_x^{(1)}, \dots, \tilde{b}_x^{(k)}\}$ bezüglich

$$(g, h)_x = \sum_{i=1}^n g(x_i) h(x_i) w(x - x_i), \quad g, h \in C(K; \mathbb{R}),$$

wobei die Funktionen $\tilde{b}_x^{(i)}$ gemäß (1.23) und (1.24) gebildet werden. Dann wird die Funktion $\text{LS-}\mathcal{L}_x f$ als *lokal Interpolierende von f in x im Sinne von Lancaster und Šalkauskas* bezeichnet.

Die lokal Interpolierende $\text{LS-}\mathcal{L}_x f$ nach Definition 1.6 lässt sich durch

$$\text{LS-}\mathcal{L}_x f = \tilde{a}_1(x) \tilde{b}_x^{(1)} + \sum_{i=2}^k \tilde{a}_i(x) \tilde{b}_x^{(i)} \quad (1.27)$$

für gewisse Koeffizienten $\tilde{a}_1(x), \dots, \tilde{a}_k(x) \in \mathbb{R}$ darstellen. Wir zeigen im weiteren Verlauf, dass die Koeffizienten auch für $x \in \{x_1, \dots, x_n\}$ existieren, insbesondere ist damit die lokal Interpolierende im Sinne von Lancaster und Šalkauskas wohldefiniert.

Für $x \in K \setminus \{x_1, \dots, x_n\}$ berechnen sich die Koeffizienten $\tilde{a}_i(x)$ durch Lösung der zum Bestapproximationsproblem gehörenden Normalengleichungen

$$\tilde{a}_1(x) (\tilde{b}_x^{(1)}, \tilde{b}_x^{(1)})_x + \sum_{i=2}^k \tilde{a}_i(x) (\tilde{b}_x^{(i)}, \tilde{b}_x^{(1)})_x = (f, \tilde{b}_x^{(1)})_x, \quad (1.28a)$$

$$\tilde{a}_1(x) (\tilde{b}_x^{(1)}, \tilde{b}_x^{(j)})_x + \sum_{i=2}^k \tilde{a}_i(x) (\tilde{b}_x^{(i)}, \tilde{b}_x^{(j)})_x = (f, \tilde{b}_x^{(j)})_x, \quad 2 \leq j \leq k. \quad (1.28b)$$

Wird in Gleichung (1.28a) der Koeffizient $\tilde{a}_1(x)$ für beliebiges $x \in K$ fest durch

$$\tilde{a}_1(x) = (f, \tilde{b}_x^{(1)})_x,$$

gewählt, ist zum einen

$$\tilde{a}_1(x) \tilde{b}_x^{(1)} = (f, \tilde{b}_x^{(1)})_x \tilde{b}_x^{(1)} = \mathcal{S}_n[w] f(x),$$

womit insbesondere eine natürliche Hebung der Singularität $\lim_{x \rightarrow \tilde{x}} \tilde{b}_x^{(1)} = +\infty$ für $\tilde{x} \in \{x_1, \dots, x_n\}$ gegeben ist, und zum anderen berechnen sich die restlichen Koeffizienten durch Lösung eines reduzierten Gleichungssystems

$$\sum_{i=2}^k \tilde{a}_i(x) (\tilde{b}_x^{(i)}, \tilde{b}_x^{(j)})_x = (f - \mathcal{S}_n[w] f(x), \tilde{b}_x^{(j)})_x, \quad 2 \leq j \leq k, \quad (1.29)$$

nach entsprechender Umformung von (1.28b). Für die Gleichungen (1.29) können wir zeigen, dass diese auch im Fall $\tilde{x} \in \{x_1, \dots, x_n\}$ wohldefiniert sind. Dazu genügt es die Existenz der Grenzwerte

$$\lim_{x \rightarrow \tilde{x}} (\tilde{b}_x^{(i)}, \tilde{b}_x^{(j)})_x \quad (1.30)$$

und

$$\lim_{x \rightarrow \tilde{x}} (f - \mathcal{S}_n[w] f(x), \tilde{b}_x^{(j)})_x \quad (1.31)$$

für beliebiges $\tilde{x} \in \{x_1, \dots, x_n\}$ nachzuweisen. Betrachten wir zunächst (1.30). Wiederum genügt es vermöge

$$(\tilde{b}_x^{(i)}, \tilde{b}_x^{(j)})_x = \sum_{l=1}^n \left(b^{(i)}(x_l) - \mathcal{S}_n[w]b^{(i)}(x) \right) \left(b^{(j)}(x_l) - \mathcal{S}_n[w]b^{(j)}(x) \right) w(x - x_l)$$

die Existenz des Grenzwertes

$$\lim_{x \rightarrow \tilde{x}} \left(b^{(i)}(x_l) - \mathcal{S}_n[w]b^{(i)}(x) \right) \left(b^{(j)}(x_l) - \mathcal{S}_n[w]b^{(j)}(x) \right) w(x - x_l)$$

zu beweisen. Diese folgt aufgrund von

$$\begin{aligned} & \left(b^{(i)}(x_l) - \mathcal{S}_n[w]b^{(i)}(x) \right) \left(b^{(j)}(x_l) - \mathcal{S}_n[w]b^{(j)}(x) \right) w(x - x_l) \\ &= \left(b^{(i)}(x_l) - \mathcal{S}_n[w]b^{(i)}(x) \right) \left(\sum_{\substack{m=1 \\ m \neq l}}^n (b^{(j)}(x_l) - b^{(j)}(x_m)) \frac{w(x - x_m)}{\sum_{s=1}^n w(x - x_s)} \right) w(x - x_l) \\ &= \left(b^{(i)}(x_l) - \mathcal{S}_n[w]b^{(i)}(x) \right) \left(\sum_{\substack{m=1 \\ m \neq l}}^n (b^{(j)}(x_l) - b^{(j)}(x_m)) w(x - x_m) \right) \frac{w(x - x_l)}{\sum_{s=1}^n w(x - x_s)} \end{aligned} \quad (1.32)$$

und

$$\lim_{x \rightarrow x_m} w(x - x_m) \frac{w(x - x_l)}{\sum_{s=1}^n w(x - x_s)} \in \mathbb{R} \quad (1.33)$$

für alle $m \neq l$ nach Definition 1.4. Der Grenzwert (1.31) ist analog zu behandeln: Am Ende der zugehörigen Schlussfolgerungen steht in diesem Fall der Term

$$\left(f(x_l) - \mathcal{S}_n[w]f(x) \right) \left(\sum_{\substack{m=1 \\ m \neq l}}^n (b^{(j)}(x_l) - b^{(j)}(x_m)) w(x - x_m) \right) \frac{w(x - x_l)}{\sum_{s=1}^n w(x - x_s)}.$$

Dieser Wert existiert ebenfalls für $x \rightarrow \tilde{x}$ aufgrund von (1.33).

Lemma 1.2 Sei $f \in C(K; \mathbb{R})$, $w \in C^l(\mathbb{R}^d \setminus \{0\}; \mathbb{R}_{>0})$ ein positives Interpolationsgewicht und $\{b^{(1)} \dots, b^{(k)}\}$ eine Interpolationsbasis. Für fest gewähltes $x \in K$ seien die Funktionen $\tilde{b}_x^{(1)}, \dots, \tilde{b}_x^{(k)}$ durch (1.23) und (1.24) definiert. Dann besitzt das reduzierte Gleichungssystem (1.29) eine eindeutig bestimmte Lösung $(\tilde{a}_2(x), \dots, \tilde{a}_k(x))^T \in \mathbb{R}^{k-1}$.

Beweis. Nach den vorausgehenden Betrachtungen ist nur die Regularität der Gramschen Matrix

$$\tilde{B}_x = \left((\tilde{b}_x^{(i)}, \tilde{b}_x^{(j)})_x \right)_{i,j=2,\dots,k}$$

zu beweisen. Diese folgt unmittelbar aus der linearen Unabhängigkeit der $b^{(2)}, \dots, b^{(k)}$ nach Voraussetzung und der Definition der Funktionen $\tilde{b}_x^{(2)}, \dots, \tilde{b}_x^{(k)}$. \square

Satz 1.5 Zu $f \in C(K; \mathbb{R})$ und $x \in K$ ist jede lokal Interpolierende LS- $\mathcal{L}_x f$ von f in $x \in K$ bezüglich $(\cdot, \cdot)_x$ im Sinne von Lancaster und Šalkauskas wohldefiniert und es gilt

$$\text{LS-}\mathcal{L}_x f(y) = \mathcal{S}_n[w]f(x) + \sum_{i=2}^k \tilde{a}_i(x) \tilde{b}_x^{(i)}(y), \quad y \in K,$$

wobei die Koeffizienten $(\tilde{a}_2(x), \dots, \tilde{a}_k(x))^T \in \mathbb{R}^{k-1}$ die Gleichungen (1.29) erfüllen.

Beweis. Lemma 1.2. \square

Definition 1.7 Sei $f \in C(K; \mathbb{R})$ und $\text{LS-}\mathcal{L}_x f$ die lokal Interpolierende von f in $x \in K$ unter den Voraussetzungen von Satz 1.5. Dann heit die Funktion $\text{LS-}\mathcal{G}f : K \rightarrow \mathbb{R}$, punktweise definiert durch

$$\begin{aligned} \text{LS-}\mathcal{G}f(x) &= \text{LS-}\mathcal{L}_x f(x) \\ &= \mathcal{S}_n[w]f(x) + \sum_{i=2}^k \tilde{a}_i(x) \left(b^{(i)}(x) - \mathcal{S}_n[w]b^{(i)}(x) \right), \quad x \in K, \end{aligned} \quad (1.34)$$

die *globale kleinste Quadrate Interpolierende* zu f im Sinne von Lancaster und Šalkauskas.

Theorem 1.2 Sei $f \in C(K; \mathbb{R})$, $w \in C^l(\mathbb{R}^d \setminus \{0\}; \mathbb{R}_{>0})$ ein positives Interpolationsgewicht und $\{b^{(1)}, \dots, b^{(k)}\} \subset C^m(K; \mathbb{R})$ eine Interpolationsbasis. Dann ist die globale kleinste Quadrate Interpolierende zu f im Sinne von Lancaster und Šalkauskas glatt gem

$$\text{LS-}\mathcal{G}f \in C^{\min\{l, m\}}(K; \mathbb{R})$$

und es gilt

$$\text{LS-}\mathcal{G}f(\tilde{x}) = f(\tilde{x})$$

fr alle $\tilde{x} \in \{x_1, \dots, x_n\}$.

Beweis. Zum einen folgt die Glattheitsaussage fr $\text{LS-}\mathcal{G}f$ aufgrund von $\tilde{a}_i \in C^{\min\{l, m\}}(K; \mathbb{R})$ nach Konstruktion und $\mathcal{S}_n[w]f \in C^l(K; \mathbb{R})$ sowie $b^{(i)} \in C^m(K; \mathbb{R})$. Und zum anderen ist die Interpolationseigenschaft von $\text{LS-}\mathcal{G}f$ eine Konsequenz derselben fr die Shepard-Interpolation, denn $b^{(i)}(\tilde{x}) - \mathcal{S}_n[w]b^{(i)}(\tilde{x}) = 0$ und $\mathcal{S}_n[w]f(\tilde{x}) = f(\tilde{x})$ gelten fr alle $\tilde{x} \in \{x_1, \dots, x_n\}$. \square

In der englischsprachigen Literatur wird der Ansatz nach Lancaster und Šalkauskas als *moving least squares interpolation* bezeichnet.

Im nchsten Abschnitt bringen einen auf Sonar et al. [28] basierenden verwandten Ansatz der *moving least squares interpolation*.

1.4.iii Interpolierende kleinste Quadrate: Konstruktion nach Sonar

Im Gegensatz zur Methode nach Lancaster und Šalkauskas betrachten wir an dieser Stelle analog zu (1.20) fr den Fall $\varepsilon > 0$ die Abbildung $(\cdot, \cdot)_{x, \varepsilon} : C(K; \mathbb{R}) \times C(K; \mathbb{R}) \rightarrow \mathbb{R}$ mit

$$(g, h)_{x, \varepsilon} = \sum_{i=1}^n g(x_i) h(x_i) w(x - x_i + \varepsilon), \quad g, h \in C(K; \mathbb{R}), \quad (1.35)$$

fr ein positives Interpolationsgewicht $w \in C^l(\mathbb{R}^d \setminus \{0\}; \mathbb{R}_{>0})$. Hierbei definiert (1.35) analog zu Lemma 1.1 ein inneres Produkt auf $C(K; \mathbb{R})$, denn $(g, h)_{x, \varepsilon} \in \mathbb{R}$ fr alle $g, h \in C(K; \mathbb{R})$ und $(\cdot, \cdot)_{x, \varepsilon}$ ist eine positiv definite und symmetrische Bilinearform.

An den Begriffen *lokal Interpolierende* und *globale kleinste Quadrate Interpolierende* des letztes Abschnitts orientierend geben wir vorab einige Definitionen.

Definition 1.8 Sei $f \in C(K; \mathbb{R})$ eine stetige Funktion, $\{b^{(1)}, \dots, b^{(k)}\}$ eine Interpolationsbasis und $w \in C^l(\mathbb{R}^d \setminus \{0\}; \mathbb{R}_{>0})$ ein positives Interpolationsgewicht. Zu fest gewähltem $x \in K$ und zu $\varepsilon > 0$ bezeichnet

$$\mathcal{L}_{x,\varepsilon}f = \sum_{i=1}^k a_{i,\varepsilon}(x)b^{(i)} \quad (1.36)$$

die Bestapproximation von f in $\text{span}\{b^{(1)}, \dots, b^{(k)}\}$ bezüglich des inneren Produktes

$$(g, h)_{x,\varepsilon} = \sum_{i=1}^n g(x_i)h(x_i)w(x - x_i + \varepsilon), \quad g, h \in C(K; \mathbb{R}). \quad (1.37)$$

Existieren die Grenzwerte

$$a_i(x) = \lim_{\varepsilon \rightarrow 0} a_{i,\varepsilon}(x), \quad 1 \leq i \leq k,$$

so heißt die Funktion $\mathcal{L}_x : K \rightarrow \mathbb{R}$ mit

$$\mathcal{L}_x f = \sum_{i=1}^k a_i(x)b^{(i)}$$

die *lokal Interpolierende* von f in $x \in K$. Existiert darüber hinaus $\mathcal{L}_x f$ für jedes $x \in K$, definieren wir $\mathcal{G}f : K \rightarrow \mathbb{R}$ punktweise durch

$$\mathcal{G}f(x) = \mathcal{L}_x f(x) = \sum_{i=1}^k a_i(x)b^{(i)}(x), \quad x \in K. \quad (1.38)$$

Wir bezeichnen die Funktion $\mathcal{G}f$ als *globale kleinste Quadrate Interpolierende* zu f oder auch *exakte kleinste Quadrate Interpolierende* zu f .

Der Prozess (1.38) wird in der englischsprachigen Literatur als *moving* bezeichnet.

Im Folgenden müssen wir die Wohldefiniertheit der in Definition 1.8 eingeführten Funktionen $\mathcal{L}_{x,\varepsilon}$, \mathcal{L}_x sowie $\mathcal{G}f$ beweisen. Dazu bemerken wir, dass die Koeffizienten $a_{1,\varepsilon}(x), \dots, a_{k,\varepsilon}(x)$ für fest gewähltes $x \in K$ durch Lösung der zum Bestapproximationsproblem korrespondierenden Normalgleichungen

$$\sum_{i=1}^k a_{i,\varepsilon}(x) (b^{(i)}, b^{(j)})_{x,\varepsilon} = (f, b^{(j)})_{x,\varepsilon}, \quad 1 \leq j \leq k, \quad (1.39)$$

zu berechnen sind. Hierbei schreibt sich (1.39) in Matrix-Vektor-Notation als

$$BW(x + \varepsilon)B^T \underline{a}_\varepsilon(x) = BW(x + \varepsilon)f, \quad (1.40)$$

mit der reellen $k \times n$ Matrix

$$B = \begin{pmatrix} b^{(1)}(x_1) & b^{(1)}(x_2) & \dots & b^{(1)}(x_{n-1}) & b^{(1)}(x_n) \\ b^{(2)}(x_1) & b^{(2)}(x_2) & \dots & b^{(2)}(x_{n-1}) & b^{(2)}(x_n) \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ b^{(k-1)}(x_1) & b^{(k-1)}(x_2) & \dots & b^{(k-1)}(x_{n-1}) & b^{(k-1)}(x_n) \\ b^{(k)}(x_1) & b^{(k)}(x_2) & \dots & b^{(k)}(x_{n-1}) & b^{(k)}(x_n) \end{pmatrix}, \quad (1.41)$$

der Diagonalmatrix $W(x) = \text{diag}(w(x - x_1), \dots, w(x - x_n)) \in \mathbb{R}^{n \times n}$, sowie den Vektoren $\underline{a}_\varepsilon(x) = (a_{1,\varepsilon}(x), \dots, a_{k,\varepsilon}(x))^T \in \mathbb{R}^k$ und $\underline{f} = (f(x_1), \dots, f(x_n))^T \in \mathbb{R}^n$.

Die Matrix $BW(x + \varepsilon)B^T$ ist als Gramsche Matrix des inneren Produktes $(\cdot, \cdot)_{x,\varepsilon}$ aufgrund der linearen Unabhängigkeit der $b^{(i)}$ regulär, das heißt die Koeffizienten $a_{i,\varepsilon}(x)$ berechnen sich durch

$$\underline{a}_\varepsilon(x) = \left(BW(x + \varepsilon)B^T \right)^{-1} BW(x + \varepsilon)\underline{f}, \quad x \in K. \quad (1.42)$$

Um die Frage zu beantworten wie sich (1.42) im Grenzfall $\varepsilon \rightarrow 0$ für beliebiges $x \in K$ verhält, bringen wir ein Resultat zur Invertierbarkeit von $BW(x + \varepsilon)B^T$, wie ursprünglich von Sonar et al. in [28] bewiesen. Hierfür geben wir zunächst die Sherman-Morrison-Formel an, deren Beweis bei [10] zu finden ist.

Lemma 1.3 (Sherman-Morrison-Formel) Sei $A \in \mathbb{R}^{m \times m}$ eine reguläre Matrix und $u, v \in \mathbb{R}^m$. Dann ist auch die Matrix $A + uv^T$ regulär und es gilt

$$\left(A + uv^T \right)^{-1} = A^{-1} - \frac{(A^{-1}u)(A^{-1}v)^T}{1 + v^T A^{-1}u}.$$

Nun zum angekündigten Invertierbarkeitsresultat für die Matrix $BW(x + \varepsilon)B^T$.

Lemma 1.4 Für jedes $x \in K$ existiert der Grenzwert

$$\left(BW(x)B^T \right)^{-1} := \lim_{\varepsilon \rightarrow 0} \left(BW(x + \varepsilon)B^T \right)^{-1},$$

mit $B \in \mathbb{R}^{k \times n}$ und $W(x + \varepsilon) \in \mathbb{R}^{n \times n}$ wie in (1.40). Insbesondere gilt für $x_j \in \{x_1, \dots, x_n\}$

$$\left(BW(x_j)B^T \right)^{-1} = R(x_j)^{-1} - \frac{1}{\underline{b}(x_j)^T R(x_j)^{-1} \underline{b}(x_j)} \left(R(x_j)^{-1} \underline{b}(x_j) \right) \left(R(x_j)^{-1} \underline{b}(x_j) \right)^T,$$

mit $R : \{x_1, \dots, x_n\} + \mathbb{R}^d \rightarrow \mathbb{R}^{k \times k}$ definiert für $h \in \mathbb{R}^d$ durch

$$R(x_j + h) = \begin{pmatrix} \sum_{\substack{i=1 \\ i \neq j}}^n b^{(1)}(x_i) w^{(i)}(x_j + h) b^{(1)}(x_i) & \dots & \sum_{\substack{i=1 \\ i \neq j}}^n b^{(1)}(x_i) w^{(i)}(x_j + h) b^{(k)}(x_i) \\ \vdots & \ddots & \vdots \\ \sum_{\substack{i=1 \\ i \neq j}}^n b^{(k)}(x_i) w^{(i)}(x_j + h) b^{(1)}(x_i) & \dots & \sum_{\substack{i=1 \\ i \neq j}}^n b^{(k)}(x_i) w^{(i)}(x_j + h) b^{(k)}(x_i) \end{pmatrix} \quad (1.43)$$

und $\underline{b}(x) = (b^{(1)}(x), \dots, b^{(k)}(x))^T \in \mathbb{R}^k$ für $x \in \mathbb{R}^d$.

Beweis. Für $x_j \in \{x_1, \dots, x_n\}$ ist $BW(x_j + \varepsilon)B^T = R(x_j + \varepsilon) + w(\varepsilon)\underline{b}(x_j)\underline{b}(x_j)^T$. Aus der Sherman-Morrison-Formel folgt

$$\begin{aligned} & \left(BW(x_j + \varepsilon)B^T \right)^{-1} \\ &= R(x_j + \varepsilon)^{-1} - \frac{w(\varepsilon)}{1 + w(\varepsilon) \sigma_\varepsilon(x_j, x_j)} \left(R(x_j + \varepsilon)^{-1} \underline{b}(x_j) \right) \left(R(x_j + \varepsilon)^{-1} \underline{b}(x_j) \right)^T \\ &= R(x_j + \varepsilon)^{-1} - \frac{1}{w(\varepsilon)^{-1} + \sigma_\varepsilon(x_j, x_j)} \left(R(x_j + \varepsilon)^{-1} \underline{b}(x_j) \right) \left(R(x_j + \varepsilon)^{-1} \underline{b}(x_j) \right)^T, \end{aligned}$$

wobei wir $\sigma_\varepsilon : \{x_1, \dots, x_n\} \times \{x_1, \dots, x_n\} \rightarrow \mathbb{R}$ durch

$$\sigma_\varepsilon(x_i, x_j) := \underline{b}(x_i)^T R(x_i + \varepsilon)^{-1} \underline{b}(x_j), \quad 1 \leq i, j \leq n,$$

definiert haben. Zum einen existiert nun der Grenzwert

$$\sigma(x_j, x_j) = \lim_{\varepsilon \rightarrow 0} \sigma_\varepsilon(x_j, x_j) \quad (1.44)$$

nach Konstruktion der Abbildung R und zum anderen ist $\lim_{\varepsilon \rightarrow 0} w(\varepsilon)^{-1} = 0$. Hiermit folgt

$$\lim_{\varepsilon \rightarrow 0} \left(BW(x_j + \varepsilon) B^T \right)^{-1} = R(x_j)^{-1} - \frac{1}{\sigma(x_j, x_j)} \left(R(x_j)^{-1} \underline{b}(x_j) \right) \left(R(x_j)^{-1} \underline{b}(x_j) \right)^T.$$

□

Mit Hilfe von Lemma 1.4 können wir die Existenz von (1.42) für den Grenzfall $\varepsilon \rightarrow 0$ beweisen.

Satz 1.6 *Der Grenzwert*

$$\underline{a}(x) = \lim_{\varepsilon \rightarrow 0} \underline{a}_\varepsilon(x) \in \mathbb{R}^k$$

mit $\underline{a}_\varepsilon(x)$ gemäß (1.42) existiert für jedes $x \in K$. Insbesondere gilt für $x_j \in \{x_1, \dots, x_n\}$ die Identität

$$\underline{a}(x_j) = \sum_{\substack{i=1 \\ i \neq j}}^n f_i w^{(i)}(x_j) R(x_j)^{-1} \left(\underline{b}(x_i) - \frac{\sigma(x_j, x_i)}{\sigma(x_j, x_j)} \underline{b}(x_j) \right) + \frac{f_j}{\sigma(x_j, x_j)} R(x_j)^{-1} \underline{b}(x_j), \quad (1.45)$$

mit $w^{(i)} \equiv w(\cdot - x_i)$, $f_i \equiv f(x_i)$ sowie $\sigma(\cdot, \cdot)$ nach (1.44).

Beweis. Für $x \in K \setminus \{x_1, \dots, x_n\}$ ist die Existenz des Grenzwertes offensichtlich, denn $W(x)$ und $BW(x)B^T$ sind in diesem Fall reguläre Matrizen. Für $x \in \{x_1, \dots, x_n\}$ folgt mit Lemma 1.4

$$\underline{a}_\varepsilon(x_j) = \left(R(x_j + \varepsilon)^{-1} - \frac{w^{(j)}(x_j + \varepsilon)}{1 + w^{(j)}(x_j + \varepsilon) \sigma_\varepsilon(x_j, x_j)} (R(x_j + \varepsilon)^{-1} \underline{b}(x_j)) (R(x_j + \varepsilon)^{-1} \underline{b}(x_j))^T \right) BW(x_j + \varepsilon) \underline{f}.$$

Elementare Umformungen liefern

$$\underline{a}_\varepsilon(x_j) = \sum_{i=1}^n f_i w^{(i)}(x_j + \varepsilon) \left(R(x_j + \varepsilon)^{-1} \underline{b}(x_i) - \frac{w^{(j)}(x_j + \varepsilon) \sigma_\varepsilon(x_j, x_i)}{1 + w^{(j)}(x_j + \varepsilon) \sigma_\varepsilon(x_j, x_j)} R(x_j + \varepsilon)^{-1} \underline{b}(x_j) \right). \quad (1.46)$$

Betrachten wir nun voneinander getrennt die Terme mit $i = j$ und $i \neq j$ in der Summe (1.46).

Für $i = j$ gilt

$$\begin{aligned} & f_j w^{(j)}(x_j + \varepsilon) \left(R(x_j + \varepsilon)^{-1} \underline{b}(x_j) - \frac{w^{(j)}(x_j + \varepsilon) \sigma_\varepsilon(x_j, x_j)}{1 + w^{(j)}(x_j + \varepsilon) \sigma_\varepsilon(x_j, x_j)} R(x_j + \varepsilon)^{-1} \underline{b}(x_j) \right) \\ &= \frac{f_j w^{(j)}(x_j + \varepsilon)}{1 + w^{(j)}(x_j + \varepsilon) \sigma_\varepsilon(x_j, x_j)} R(x_j + \varepsilon)^{-1} \underline{b}(x_j) \\ &= \frac{f_j}{(w^{(j)}(x_j + \varepsilon))^{-1} + \sigma_\varepsilon(x_j, x_j)} R(x_j + \varepsilon)^{-1} \underline{b}(x_j) \longrightarrow \frac{f_j}{\sigma(x_j, x_j)} R(x_j)^{-1} \underline{b}(x_j) \end{aligned}$$

mit $\varepsilon \rightarrow 0$ und im Fall $i \neq j$ erhalten wir

$$\begin{aligned} f_i w^{(i)}(x_j + \varepsilon) & \left(R(x_j + \varepsilon)^{-1} \underline{b}(x_i) - \frac{w^{(j)}(x_j + \varepsilon) \sigma_\varepsilon(x_j, x_i)}{1 + w^{(j)}(x_j + \varepsilon) \sigma_\varepsilon(x_j, x_j)} R(x_j + \varepsilon)^{-1} \underline{b}(x_j) \right) \\ & = f_i w^{(i)}(x_j + \varepsilon) R(x_j + \varepsilon)^{-1} \left(\underline{b}(x_i) - \frac{\sigma_\varepsilon(x_j, x_i)}{(w^{(j)}(x_j + \varepsilon))^{-1} + \sigma_\varepsilon(x_j, x_j)} \underline{b}(x_j) \right) \\ & \longrightarrow f_i w^{(i)}(x_j) R(x_j)^{-1} \left(\underline{b}(x_i) - \frac{\sigma(x_j, x_i)}{\sigma(x_j, x_j)} \underline{b}(x_j) \right) \end{aligned}$$

mit $\varepsilon \rightarrow 0$. Dies liefert die Behauptung. \square

Mit $\underline{a}(x) = (a_1(x), \dots, a_k(x))^T \in \mathbb{R}^k$ liefert uns Satz 1.6 sowohl die Wohldefiniertheit der lokal Interpolierenden $\mathcal{L}_x f$ der Funktion $f \in C(K; \mathbb{R})$ im Punkt $x \in K$, als auch die der daraus bestimmten exakte kleinste Quadrate Interpolierenden $\mathcal{G}f$ nach Definition 1.8.

Theorem 1.3 Sei $f \in C(K; \mathbb{R})$, $\{b^{(1)}, \dots, b^{(k)}\} \subset C^m(K; \mathbb{R})$ eine Interpolationsbasis sowie $w \in C^l(\mathbb{R}^d \setminus \{0\}; \mathbb{R}_{>0})$ ein positives Interpolationsgewicht. Dann ist die globale bzw. exakte kleinste Quadrate Interpolierende $\mathcal{G}f$ zu f im Sinne von Sonar glatt gemäß

$$\mathcal{G}f \in C^{\min\{l, m\}}(K; \mathbb{R})$$

und es gilt

$$\mathcal{G}f(x_j) = f(x_j)$$

für alle $x_j \in \{x_1, \dots, x_n\}$.

Beweis. Die Differenzierbarkeitsaussage für $\mathcal{G}f$ folgt aufgrund von $a_i \in C^l(K; \mathbb{R})$ nach deren Konstruktion sowie (1.42) und $b^{(i)} \in C^m(K; \mathbb{R})$. Zum Beweis der Interpolationseigenschaft ist $\underline{a}(x_j)^T \underline{b}(x_j) = f(x_j)$ für alle $x_j \in \{x_1, \dots, x_n\}$ und $\underline{a}(x)$ nach Satz 1.6 zu zeigen. Für $x = x_j$ und $i \neq j$ gilt

$$\begin{aligned} f_i w^{(i)}(x_j) & \left(R(x_j)^{-1} \left(\underline{b}(x_i) - \frac{\sigma(x_j, x_i)}{\sigma(x_j, x_j)} \underline{b}(x_j) \right) \right)^T \underline{b}(x_j) \\ & = f_i w^{(i)}(x_j) \left(\sigma(x_j, x_i) - \frac{\sigma(x_j, x_i)}{\sigma(x_j, x_j)} \sigma(x_j, x_j) \right) = 0 \end{aligned}$$

sowie

$$\frac{f_j}{\sigma(x_j, x_j)} \left(R(x_j)^{-1} \underline{b}(x_j) \right)^T \underline{b}(x_j) = f_j,$$

nach Satz 1.6. \square

An dieser Stelle notieren wir noch kurz, dass ähnlich wie im Fall der Shepard-Interpolation die Koeffizienten $a_{i,\varepsilon}(x)$ in (1.36) durch Minimierung des quadratischen Funktionals

$$Q_{x,\varepsilon}(a_{1,\varepsilon}(x), \dots, a_{k,\varepsilon}(x)) = \sum_{j=1}^n \left(\sum_{i=1}^k a_{i,\varepsilon}(x) b^{(i)}(x_j) - f(x_j) \right)^2 w(x - x_j + \varepsilon)$$

gewonnen werden können.

In diesem Abschnitt haben wir zwei Ansätze der kleinste Quadrate Interpolation kennengelernt. Zum einen die Methode nach Lancaster und Šalkauskas und zum anderen die exakte kleinste Quadrate Interpolation nach Sonar. Beide Methoden beruhen im Wesentlichen auf ein und derselben Bestapproximationsidee, jedoch werden wir im folgenden Abschnitt, welcher einen rigorosen Vergleich der beiden Ansätze liefert.

1.5 Ein Vergleich von $\text{LS-}\mathcal{G}f$ und $\mathcal{G}f$

Zu $f \in C(K; \mathbb{R})$, fest gewählten $\{x_1, \dots, x_n\} \subset K$ mit $x_i \neq x_j$ für $i \neq j$, einer Interpolationsbasis $\{b^{(1)}, \dots, b^{(k)}\} \subset C^m(K; \mathbb{R})$ und einem positiven Interpolationsgewicht $w \in C^l(\mathbb{R}^d \setminus \{0\}; \mathbb{R}_{>0})$ bezeichne $\text{LS-}\mathcal{G}f$ die globale kleinste Quadrate Interpolierende zu f im Sinne von Lancaster und Šalkauskas nach Definition 1.7 sowie $\mathcal{G}f$ die entsprechende exakte kleinste Quadrate Interpolierende zu f nach Definition 1.8 mit

$$\text{LS-}\mathcal{G}f(x) = \mathcal{S}_n[w]f(x) + \sum_{i=2}^k \tilde{a}_i(x) \left(b^{(i)}(x) - \mathcal{S}_n[w]b^{(i)}(x) \right) \quad (1.47)$$

und

$$\mathcal{G}f(x) = \sum_{i=1}^k a_i(x) b^{(i)}(x), \quad (1.48)$$

für Koeffizienten $\tilde{a}_i(x)$ und $a_i(x)$ entsprechend den Abschnitten 1.4.ii sowie 1.4.iii. Offensichtlich ist die Darstellung (1.48) äquivalent zu

$$\mathcal{G}f(x) = \sum_{i=1}^k a_i(x) \mathcal{S}_n[w]b^{(i)}(x) + \sum_{i=2}^k a_i(x) \left(b^{(i)}(x) - \mathcal{S}_n[w]b^{(i)}(x) \right), \quad (1.49)$$

unter Beachtung von $b^{(1)} = \mathcal{S}_n[w]b^{(1)}$, wodurch insbesondere ein direkter Vergleich von (1.49) mit (1.47) auf globaler Ebene möglich wird.

Satz 1.7 Seien $\text{LS-}\mathcal{G}f$ und $\mathcal{G}f$ wie in (1.47) und (1.48). Gilt für $x \in K$ die Identität

$$\mathcal{S}_n[w]f(x) = \sum_{i=1}^k a_i(x) \mathcal{S}_n[w]b^{(i)}(x), \quad (1.50)$$

sowie $\tilde{a}_i(x) = a_i(x)$ für $2 \leq i \leq k$, so ist $\text{LS-}\mathcal{G}f(x) = \mathcal{G}f(x)$.

Beweis. Die Behauptung folgt durch Vergleich von (1.47) und (1.48) bzw. (1.49). \square

Bei näherer Betrachtung von (1.50) erkennen wir sofort, dass diese Identität trivialerweise für jedes $\tilde{x} \in \{x_1, \dots, x_n\}$ aufgrund von

$$\mathcal{S}_n[w]f(\tilde{x}) = f(\tilde{x}) = \mathcal{G}f(\tilde{x}) = \sum_{i=1}^k a_i(\tilde{x}) b^{(i)}(\tilde{x}) = \sum_{i=1}^k a_i(\tilde{x}) \mathcal{S}_n[w]b^{(i)}(\tilde{x}) \quad (1.51)$$

erfüllt ist. Würde ferner zu $\tilde{x} \in \{x_1, \dots, x_n\}$ ein $i \in \{2, \dots, k\}$ mit $\tilde{a}_i(\tilde{x}) \neq a_i(\tilde{x})$ existieren, stände dies im Widerspruch zu $\text{LS-}\mathcal{G}f(\tilde{x}) = \mathcal{G}f(\tilde{x})$ aufgrund von (1.51), das heißt insbesondere ist $\tilde{a}_i(\tilde{x}) = a_i(\tilde{x})$ für $2 \leq i \leq k$ und $\tilde{x} \in \{x_1, \dots, x_n\}$. Jedoch können wir zunächst auf globaler

Ebene die Gültigkeit von (1.50) nicht für beliebiges $x \in K$ garantieren. Anstatt des Vergleichs der globalen Interpolierenden etablieren wir nun einen Vergleich auf lokaler Ebene. Hier stellen wir sofort fest, dass für jedes $x \in K \setminus \{x_1, \dots, x_n\}$ die Funktionen $\text{LS-}\mathcal{L}_x f$ und $\mathcal{L}_x f$ aufgrund von

$$\text{span}\{\tilde{b}_x^{(1)}, \dots, \tilde{b}_x^{(k)}\} = \text{span}\{b^{(1)}, \dots, b^{(k)}\}$$

identisch sind. Für $\tilde{x} \in \{x_1, \dots, x_n\}$ haben wir aber bereits $\text{LS-}\mathcal{L}_{\tilde{x}} f(\tilde{x}) = \mathcal{L}_{\tilde{x}} f(\tilde{x})$ festgestellt. Insgesamt resultiert

$$\text{LS-}\mathcal{L}_x f(x) = \mathcal{L}_x f(x)$$

für beliebiges $x \in K$.

Wir hatten in den vorausgehenden Abschnitten festgestellt, dass sich die Glattheit der Interpolationsbasisfunktionen sowie die des positiven Interpolationsgewichtes in gewisser Weise auf die kleinste Quadrate Interpolierende überträgt. In Kapitel 2 werden wir der Frage nachgehen, wie exakte kleinste Quadrate Interpolierende nach Abschnitt 1.4.iii im Rahmen einer Approximation von Ableitungswerten genutzt werden können. In diesem Kontext stellt sich beispielsweise auf globaler Ebene die Frage nach dem Übereinstimmen von Ableitungen erster Ordnung der Interpolierenden $\text{LS-}\mathcal{G}f$ und $\mathcal{G}f$ an den Interpolationsknoten $\tilde{x} \in \{x_1, \dots, x_n\}$.

Formal berechnen sich partielle Ableitungen der Interpolierenden für $1 \leq j \leq d$ gemäß

$$\frac{\partial}{\partial x_j} \text{LS-}\mathcal{G}f(x) = \frac{\partial}{\partial x_j} \mathcal{S}f(x) + \sum_{i=2}^k \left(\frac{\partial}{\partial x_j} \tilde{a}_i(x) \mathcal{S}b^{(i)}(x) + \tilde{a}_i(x) \frac{\partial}{\partial x_j} \mathcal{S}b^{(i)}(x) \right), \quad (1.52)$$

und

$$\frac{\partial}{\partial x_j} \mathcal{G}f(x) = \sum_{i=1}^n \left(\frac{\partial}{\partial x_j} a_i(x) b^{(i)}(x) + a_i(x) \frac{\partial}{\partial x_j} b^{(i)}(x) \right), \quad (1.53)$$

wodurch deutlich wird, dass ein direkter Vergleich von (1.52) mit (1.53) im Allgemeinen nicht von Erfolg gekrönt sein wird. Hingegen ist für $\tilde{x} \in \{x_1, \dots, x_n\}$ aufgrund des *flat spot*-Phänomens des verallgemeinerten Shepard-Interpolationsoperators

$$\frac{\partial}{\partial x_j} \text{LS-}\mathcal{G}f(\tilde{x}) = \sum_{i=2}^k \frac{\partial}{\partial x_j} \tilde{a}_i(\tilde{x}) b^{(i)}(\tilde{x}), \quad (1.54)$$

so dass wir ohne Mühe ein Analogon zu Satz 1.7 für Ableitungswerte formulieren können.

Satz 1.8 Seien $\text{LS-}\mathcal{G}f$ und $\mathcal{G}f$ wie in (1.47) und (1.48) mindestens $C^1(K; \mathbb{R})$ -Funktionen, das heißt $\min\{l, m\} \geq 1$. Gilt für $\tilde{x} \in \{x_1, \dots, x_n\}$

$$\frac{\partial}{\partial x_j} a_1(\tilde{x}) b^{(1)}(\tilde{x}) + \sum_{i=1}^k a_i(\tilde{x}) \frac{\partial}{\partial x_j} b^{(i)}(\tilde{x}) = 0$$

und $\frac{\partial}{\partial x_j} \tilde{a}_i(\tilde{x}) = \frac{\partial}{\partial x_j} a_i(\tilde{x})$ für alle $2 \leq i \leq k$, dann ist $\frac{\partial}{\partial x_j} \text{LS-}\mathcal{G}f(\tilde{x}) = \frac{\partial}{\partial x_j} \mathcal{G}f(\tilde{x})$.

Beweis. Vergleich von (1.53) mit (1.54) für $\tilde{x} \in \{x_1, \dots, x_n\}$. □

Die in Satz 1.8 benötigten Ableitungswerte $\frac{\partial}{\partial x_j} a_i(\tilde{x})$ stellen wir in Satz 1.9 bereit.

Satz 1.9 Sei $\tilde{x} = x_j \in \{x_1, \dots, x_n\}$ und $\mu \in \mathbb{R}^d$ ein Multi-Index von der Ordnung $|\mu| = 1$. Für das positive Interpolationsgewicht $w \in C^1(\mathbb{R}^d \setminus \{0\}; \mathbb{R}_{>0})$ verhalte sich der Term

$$\frac{w(\varepsilon)}{D^\mu w(\varepsilon)} \quad (1.55)$$

linear in ε und es bezeichne $\eta(\tilde{x}) \in \mathbb{R} \cup \{\pm\infty\}$ den Wert

$$\eta(\tilde{x}) = \lim_{\varepsilon \rightarrow 0} \left(f(\tilde{x}) - \sum_{i=1}^k a_i(\tilde{x} + \varepsilon) b^{(i)}(\tilde{x}) \right) \frac{D^\mu w(\varepsilon)}{1 + w(\varepsilon) \sigma_\varepsilon(\tilde{x}, \tilde{x})}.$$

Sei weiter für $\varepsilon > 0$ der Wert $D^\mu \underline{a}_\varepsilon(\tilde{x})$ definiert durch

$$D^\mu \underline{a}_\varepsilon(\tilde{x}) = D^\mu \left((BW(x + \varepsilon) B^T)^{-1} BW(x + \varepsilon) \underline{f} \right) \Big|_{x=\tilde{x}}.$$

Dann existiert der Grenzwert $D^\mu \underline{a}(\tilde{x}) = \lim_{\varepsilon \rightarrow 0} D^\mu \underline{a}_\varepsilon(\tilde{x})$ und es gilt

$$\begin{aligned} D^\mu \underline{a}(\tilde{x}) &\equiv D^\mu \underline{a}(x_j) \\ &= \sum_{\substack{m=1 \\ m \neq j}}^n \left(f_m - \sum_{i=1}^k a_i(x_j) b^{(i)}(x_m) \right) \left(D^\mu w^{(m)}(x_j) \right) R(x_j)^{-1} \left(\underline{b}(x_m) - \frac{\sigma(x_j, x_m)}{\sigma(x_j, x_j)} \underline{b}(x_j) \right) \\ &\quad + \eta(x_j) R(x_j)^{-1} \underline{b}(x_j). \end{aligned}$$

Zum Beweis dieses Resultates benötigen wir das folgende Lemma.

Lemma 1.5 Sei $A = A(t) = (a_{i,j}(t))_{i,j=1}^m \in \mathbb{R}^{m \times m}$ eine reguläre Matrix mit den Koeffizientenfunktionen $a_{i,j} \in C^{(1)}([a, b]; \mathbb{R})$ für $t \in [a, b]$. Dann gilt $\frac{d}{dt} A^{-1}(t) = -A^{-1}(t) \frac{d}{dt} A(t) A^{-1}(t)$ für $t \in [a, b]$.

Beweis. Bezeichnet I_m die $(m \times m)$ -Einheitsmatrix, dann ist $0 = \frac{d}{dt} I_m = \frac{d}{dt} (A(t) A^{-1}(t))$ unter Benutzung der Produktregel zur Differentiation von Matrixprodukten, wobei durch 0 die $(m \times m)$ -Nullmatrix gegeben ist. \square

Beweis von Satz 1.9. Nach Lemma 1.5 ist

$$\begin{aligned} D^\mu \underline{a}(x_j + \varepsilon) &= D^\mu \left[\left(BW(x_j + \varepsilon) B^T \right)^{-1} \right] BW(x_j + \varepsilon) \underline{f} + \left(BW(x_j + \varepsilon) B^T \right)^{-1} D^\mu \left[BW(x_j + \varepsilon) \underline{f} \right] \\ &= - \left(BW(x_j + \varepsilon) B^T \right)^{-1} BD^\mu W(x_j + \varepsilon) B^T \underbrace{\left(BW(x_j + \varepsilon) B^T \right)^{-1} BW(x_j + \varepsilon) \underline{f}}_{=\underline{a}(x_j + \varepsilon)} \\ &\quad + \left(BW(x_j + \varepsilon) B^T \right)^{-1} BD^\mu W(x_j + \varepsilon) \underline{f} \\ &= \left(BW(x_j + \varepsilon) B^T \right)^{-1} BD^\mu W(x_j + \varepsilon) \left(\underline{f} - B^T \underline{a}(x_j + \varepsilon) \right), \end{aligned}$$

so dass unter Berücksichtigung von

$$BD^\mu W(x_j + \varepsilon) \left(\underline{f} - B^T \underline{a}(x_j + \varepsilon) \right) = \sum_{m=1}^n \left(f_m - \sum_{i=1}^k a_i(x_j + \varepsilon) b^{(i)}(x_m) \right) (D^\mu w^{(m)}(x_j + \varepsilon)) \underline{b}(x_m)$$

und Lemma 1.4

$$\begin{aligned} D^\mu \underline{a}(x_j + \varepsilon) &= \sum_{m=1}^n \left(f_m - \sum_{i=1}^k a_i(x_j + \varepsilon) b^{(i)}(x_m) \right) (D^\mu w^{(m)}(x_j + \varepsilon)) R(x_j + \varepsilon)^{-1} \bullet \dots \\ &\dots \bullet \left(\underline{b}(x_m) - \frac{w^{(j)}(x_j + \varepsilon) \sigma_\varepsilon(x_j, x_m)}{1 + w^{(j)}(x_j + \varepsilon) \sigma_\varepsilon(x_j, x_j)} \underline{b}(x_j) \right) \end{aligned} \quad (1.56)$$

folgt. Obwohl die Gleichungen (1.56) und (1.46) aus dem Beweis zu Satz 1.6 eine ähnliche Struktur besitzen, erweist sich die Analyse der Terme für den Grenzfall $\varepsilon \rightarrow 0$ als komplizierter. Betrachten wir die Terme $m \neq j$ und $m = j$ in der Summe (1.56) getrennt voneinander.

Für $m \neq j$ ist

$$\begin{aligned} &\left(f_m - \sum_{i=1}^k a_i(x_j + \varepsilon) b^{(i)}(x_m) \right) (D^\mu w^{(m)}(x_j + \varepsilon)) R(x_j + \varepsilon)^{-1} \bullet \dots \\ &\dots \bullet \left(\underline{b}(x_m) - \frac{w(\varepsilon) \sigma_\varepsilon(x_j, x_m)}{1 + w(\varepsilon) \sigma_\varepsilon(x_j, x_j)} \underline{b}(x_j) \right) \\ &= \left(f_m - \sum_{i=1}^k a_i(x_j + \varepsilon) b^{(i)}(x_m) \right) (D^\mu w^{(m)}(x_j + \varepsilon)) R(x_j + \varepsilon)^{-1} \bullet \dots \\ &\dots \bullet \left(\underline{b}(x_m) - \frac{\sigma_\varepsilon(x_j, x_m)}{1/w(\varepsilon) + \sigma_\varepsilon(x_j, x_j)} \underline{b}(x_j) \right) \\ &\longrightarrow \left(f_m - \sum_{i=1}^k a_i(x_j) b^{(i)}(x_m) \right) (D^\mu w^{(m)}(x_j)) R(x_j)^{-1} \left(\underline{b}(x_m) - \frac{\sigma(x_j, x_m)}{\sigma(x_j, x_j)} \underline{b}(x_j) \right) \end{aligned}$$

mit $\varepsilon \rightarrow 0$ und für $m = j$ ist

$$\begin{aligned} &\left(f_j - \sum_{i=1}^k a_i(x_j + \varepsilon) b^{(i)}(x_j) \right) (D^\mu w(\varepsilon)) R(x_j + \varepsilon)^{-1} \left(\underline{b}(x_j) - \frac{w(\varepsilon) \sigma_\varepsilon(x_j, x_j)}{1 + w(\varepsilon) \sigma_\varepsilon(x_j, x_j)} \underline{b}(x_j) \right) \\ &= \left(f_j - \sum_{i=1}^k a_i(x_j + \varepsilon) b^{(i)}(x_j) \right) \frac{D^\mu w(\varepsilon)}{1 + w(\varepsilon) \sigma_\varepsilon(x_j, x_j)} R(x_j + \varepsilon)^{-1} \underline{b}(x_j). \end{aligned} \quad (1.57)$$

Um die Existenz von (1.57) im Grenzfall $\varepsilon \rightarrow 0$ zu gewährleisten, sind an dieser Stelle die Voraussetzungen des Satzes bezüglich der Interpolationsbasisfunktionen und des Interpolationsgewichtes anzuwenden, das heißt aus der Interpolationseigenschaft

$$\lim_{\varepsilon \rightarrow 0} \left(f_j - \sum_{i=1}^k a_i(x_j + \varepsilon) b^{(i)}(x_j) \right) = 0,$$

folgt aufgrund von $a_i \in C(K; \mathbb{R})$ schon

$$f_j - \sum_{i=1}^k a_i(x_j + \varepsilon) b^{(i)}(x_j) = \sum_{i=1}^k (a_i(x_j) - a_i(x_j + \varepsilon)) b^{(i)}(x_j) = \mathcal{O}(\varepsilon).$$

Ferner gilt

$$\left| \frac{D^\mu w(\varepsilon)}{1 + w(\varepsilon) \sigma_\varepsilon(x_j, x_j)} \right| \leq \left| \frac{1}{\sigma_\varepsilon(x_j, x_j)} \frac{D^\mu w(\varepsilon)}{w(\varepsilon)} \right| = \text{const} \cdot \frac{1}{\varepsilon}$$

unter Berücksichtigung der Existenz des Grenzwertes $\lim_{\varepsilon \rightarrow 0} \sigma_\varepsilon(x_j, x_j)$ und $\sigma_\varepsilon(x_j, x_j) \geq 0$. Hiermit folgt sofort die Existenz von

$$\eta(\tilde{x}) \equiv \eta(x_j) = \lim_{\varepsilon \rightarrow 0} \left(f_j - \sum_{i=1}^k a_i(x_j + \varepsilon) b^{(i)}(x_j) \right) \frac{D^\mu w(\varepsilon)}{1 + w(\varepsilon) \sigma_\varepsilon(x_j, x_j)}.$$

□

1.6 Eine Kardinalbasis-Darstellung der exakte kleinste Quadrate Interpolierenden

In diesem Abschnitt bringen wir eine Darstellung der exakte kleinste Quadrate Interpolierenden in Analogie zur Darstellung der Interpolierenden nach Shepard aus Definition 1.2 bzw. Gleichung (1.25). Hierbei besteht die Aufgabe in der Gleichung

$$\underline{a}(x)^T \underline{b}(x) = \underline{f}^T \underline{\varphi}(x), \quad (1.58)$$

wobei $\underline{a}(x) = (a_1(x), \dots, a_k(x))^T$ entsprechend der Definition 1.8 zu berechnen ist, den Vektor der so genannten *Kernfunktionen*

$$\underline{\varphi}(x) = (\varphi_1(x), \dots, \varphi_n(x))^T$$

zu bestimmen und diese Funktionen zu analysieren. Aufgrund von $\underline{a}(x) = \lim_{\varepsilon \rightarrow 0} \underline{a}_\varepsilon(x)$ mit

$$\underline{a}_\varepsilon(x) = \left(BW(x + \varepsilon) B^T \right)^{-1} BW(x + \varepsilon) \underline{f}, \quad x \in K,$$

erhalten wir

$$\underline{\varphi}(x) = \lim_{\varepsilon \rightarrow 0} \left(W(x + \varepsilon) B^T \left(BW(x + \varepsilon) B^T \right)^{-1} \underline{b}(x) \right). \quad (1.59)$$

Insbesondere liefert Theorem 1.3 die Kronecker- δ -Eigenschaft

$$\varphi_i(x_j) = \delta_{ij}, \quad 1 \leq i, j \leq n. \quad (1.60)$$

Wir wollen dieses Ergebnis noch einmal explizit beweisen.

Satz 1.10 *Der Grenzwert (1.59) existiert für beliebiges $x \in K$ und es gilt*

$$\underline{\varphi}(x_j) = (\delta_{ij})_{i=1, \dots, n}^T \in \mathbb{R}^n$$

für alle $x_j \in \{x_1, \dots, x_n\}$.

Beweis. Im Fall $x \in K \setminus \{x_1, \dots, x_n\}$ ist die Existenz des Grenzwertes offensichtlich, denn $W(x)$ als auch $BW(x)B^T$ sind regulär. Für $x_j \in \{x_1, \dots, x_n\}$ und

$$\hat{\varphi}(\xi) = W(\xi) B^T \left(BW(\xi) B^T \right)^{-1} \underline{b}(\xi)$$

gilt

$$\begin{aligned} \hat{\varphi}_i(x_j + \varepsilon) &= w^{(i)}(x_j + \varepsilon) \underline{b}(x_i)^T \left(R(x_j + \varepsilon)^{-1} - \frac{w^{(j)}(x_j + \varepsilon)}{1 + w^{(j)}(x_j + \varepsilon) \sigma_\varepsilon(x_j, x_j)} \bullet \dots \right. \\ &\quad \left. \dots \bullet \left(R(x_j + \varepsilon)^{-1} \underline{b}(x_j) \right) \left(R(x_j + \varepsilon)^{-1} \underline{b}(x_j) \right)^T \right) \underline{b}(x_j + \varepsilon), \quad 1 \leq i \leq n, \end{aligned}$$

so dass aufgrund der Stetigkeit der Vektorfunktion \underline{b} in x_j schon für $i = j$

$$\begin{aligned}\hat{\varphi}_j(x_j + \varepsilon) &= w^{(j)}(x_j + \varepsilon) \left(\sigma_\varepsilon(x_j, x_j) - \frac{w^{(j)}(x_j + \varepsilon) \sigma_\varepsilon(x_j, x_j)^2}{1 + w^{(j)}(x_j + \varepsilon) \sigma_\varepsilon(x_j, x_j)} \right) + \mathcal{O}(\varepsilon) \\ &= w^{(j)}(x_j + \varepsilon) \frac{\sigma_\varepsilon(x_j, x_j)}{1 + w^{(j)}(x_j + \varepsilon) \sigma_\varepsilon(x_j, x_j)} + \mathcal{O}(\varepsilon),\end{aligned}$$

sowie für $i \neq j$

$$\hat{\varphi}_i(x_j + \varepsilon) = w^{(i)}(x_j + \varepsilon) \frac{1}{1 + w^{(j)}(x_j + \varepsilon) \sigma_\varepsilon(x_j, x_j)} \sigma_\varepsilon(x_j, x_i) + \mathcal{O}(\varepsilon)$$

gilt. Die Behauptung folgt mit $\varepsilon \rightarrow 0$. \square

Definition 1.9 Ist die exakte kleinste Quadrate Interpolierende $\mathcal{G}f \in C(K; \mathbb{R})$ zu der Funktion $f \in C(K; \mathbb{R})$ durch

$$\mathcal{G}f = \sum_{i=1}^n f(x_i) \varphi_i \quad (1.61)$$

für gewisse Kernfunktionen φ_i gegeben, dann nennen wir (1.61) *Kardinalbasis-Darstellung* der Interpolierenden. Die Kernfunktionen werden auch als *Kardinalbasisfunktionen* bezeichnet.

Ein Blick auf (1.59) zeigt, dass die Funktionen φ_i in (1.61) durch

$$\varphi_i(x) = \lim_{\varepsilon \rightarrow 0} w^{(i)}(x + \varepsilon) \underline{b}(x_i)^T \left(BW(x + \varepsilon) B^T \right)^{-1} \underline{b}(x), \quad x \in K, \quad (1.62)$$

gegeben sind.

Eine wichtige Eigenschaft des Interpolationsoperators \mathcal{G} ist, dass die unterliegenden Interpolationsbasisfunktionen invariant unter \mathcal{G} sind, das heißt diese werden durch den Operator reproduziert. Dieses Resultat kann unter Zuhilfenahme der Kardinalbasis-Darstellung der Interpolierenden sehr einfach bewiesen werden.

Satz 1.11 Es bezeichne $\mathcal{G} : C(K; \mathbb{R}) \rightarrow C(K; \mathbb{R})$ den exakte kleinste Quadrate Interpolationsoperator, der für $f \in C(K; \mathbb{R})$ durch (1.38) gegeben ist. Dann reproduziert der Operator die dessen Konstruktion unterliegenden Interpolationsbasisfunktionen $b^{(1)}, \dots, b^{(k)}$, das heißt es gilt

$$\mathcal{G}b^{(i)} = b^{(i)}, \quad 1 \leq i \leq k.$$

Beweis. Nach (1.59) gilt für die Matrix B aus (1.41) die Gleichung

$$B\underline{\varphi}(x) = \lim_{\varepsilon \rightarrow 0} \left(BW(x + \varepsilon) B^T \left(BW(x + \varepsilon) B^T \right)^{-1} \underline{b}(x) \right) = \underline{b}(x), \quad x \in K. \quad (1.63)$$

Ein komponentenweiser Vergleich von $B\underline{\varphi}(x) = \underline{b}(x)$ liefert


$$\sum_{j=1}^n b^{(i)}(x_j) \varphi_j(x) = b^{(i)}(x), \quad x \in K, \quad 1 \leq i \leq k,$$

und damit das zu beweisende Resultat. \square

Kapitel 2

Exakte kleinste Quadrate Interpolierende im Kontext von Ableitungs-Approximationen

2.1 Zur Berechnung von Ableitungen von exakte kleinste Quadrate Interpolierenden	29
2.2 Exakte kleinste Quadrate Interpolierende zur Konstruktion von Finite Differenzen	34
2.3 Inverse Euklidische Normen als positive Interpolationsgewichte	38
2.4 Die Konstruktion und Analyse symmetrischer 3-Punkte Finite Differenzen	39

ie Berechnung von Ableitungswerten von exakte kleinste Quadrate Interpolierenden ist Gegenstand des zweiten Kapitels dieser Arbeit. Hierbei werden die innerhalb der folgenden Abschnitte dargestellten Untersuchungen insbesondere im Hinblick auf die Konstruktion von Finite Differenzen-Approximationen im Kontext der numerischen Behandlung von Erhaltungsgleichungen motiviert. Wir werden die Resultate dieses Kapitels auf konkrete Fallbeispiele in Kapitel 5 anwenden.

2.1 Zur Berechnung von Ableitungen von exakte kleinste Quadrate Interpolierenden

In Theorem 1.3 haben wir gezeigt, dass die Differentiationsklasse einer exakte kleinste Quadrate Interpolierenden $\mathcal{G}f$ zu einer stetigen Funktion $f \in C(K; \mathbb{R})$ durch die Güte der zur Konstruktion von $\mathcal{G}f$ herangezogenen Interpolationsbasisfunktionen $\{b^{(1)}, \dots, b^{(k)}\} \subset C^m(K; \mathbb{R})$ sowie der des entsprechenden positiven Interpolationsgewichtes $w \in C^l(\mathbb{R}^d \setminus \{0\}; \mathbb{R}_{>0})$ bestimmt ist:

$$\mathcal{G}f \in C^{\min\{l, m\}}(K; \mathbb{R}).$$

Ist $\mu \in \mathbb{N}^d$ ein Multi-Index von der Ordnung $|\mu| \leq \min\{l, m\}$, so wird die entsprechende μ -Ableitung der Interpolierenden $\mathcal{G}f$ durch

$$D^\mu \mathcal{G}f = \sum_{i=1}^n f(x_i) D^\mu \varphi_i \tag{2.1}$$

berechnet. Bei der Konstruktion von $\mathcal{G}f$ gehen wir von einer unterliegenden Interpolationsknotenverteilung $\{x_1, \dots, x_n\} \subset K$ aus.

Aufgrund von (2.1) überträgt sich das Differentiationsproblem von der exakte kleinste Quadrate Interpolierenden $\mathcal{G}f$ unmittelbar auf das für die einzelnen Kernfunktionen φ_i . A priori stimmen allerdings die Differentiationsklassen der Kernfunktionen und die der entsprechenden exakte kleinste Quadrate Interpolierenden nicht übereinstimmen. Als einfaches Beispiel hierfür dient die Tatsache, dass auf der einen Seite $x \mapsto |x| + (-|x|)$ eine $C^\infty(-1, 1)$ -Abbildung ist und auf der anderen Seite $x \mapsto |x|$ offenkundig nicht bei $x = 0$ differenzierbar ist. Es gilt jedoch folgendes Lemma.

Lemma 2.1 *Sei $\{x_1, \dots, x_n\} \subset K$ eine Interpolationsknotenverteilung, $\{b^{(1)}, \dots, b^{(k)}\} \subset C^m(K; \mathbb{R})$ eine Interpolationsbasis und $w \in C^l(\mathbb{R}^d \setminus \{0\}; \mathbb{R}_{>0})$ ein positives Interpolationsgewicht. Dann gilt für die Kernfunktionen φ_i nach (1.62)*

$$\varphi_i \in C^{\min\{m, l\}}(K; \mathbb{R}).$$

Beweis. Die Behauptung folgt aufgrund der Darstellung

$$\varphi_i(x) = \lim_{\varepsilon \rightarrow 0} w(x - x_i + \varepsilon) \underline{b}(x_i)^T \left(BW(x + \varepsilon) B^T \right)^{-1} \underline{b}(x),$$

sowie der Konstruktion der Matrix $W = W(x)$ und des Vektors $\underline{b} = \underline{b}(x)$ durch Argumente vergleichbar mit denen im Beweis zu Theorem 1.3. \square

Eine für spätere Anwendungen interessante Aufgabe besteht in der Berechnung von Werten $D^\mu \mathcal{G}f(x_j)$ für Interpolationsknoten $x_j \in \{x_1, \dots, x_n\}$. Die Aussage von Theorem 2.1 ist für positive Interpolationsgewichte von der Form $w(x) = x^{-2\alpha}$ mit $\alpha \in \mathbb{N}$ für den univariaten Fall in [28] zu finden. Bevor wir einen von [28] unabhängigen Beweis angeben, bringen wir ein hierfür benötigtes Hilfsresultat.

Satz 2.1 *Für beliebiges $x_j \in \{x_1, \dots, x_n\}$ gilt*

$$\lim_{x \rightarrow x_j} \left(BW(x) B^T \right)^{-1} \underline{b}(x) = (0, \dots, 0)^T, \quad (2.2)$$

mit B , $W(x)$ und $\underline{b}(x)$ entsprechend Lemma 1.4.

Beweis. Die Gleichung (2.2) besitzt Gültigkeit aufgrund von

$$\begin{aligned} \left(BW(x_j + \varepsilon) B^T \right)^{-1} \underline{b}(x_j + \varepsilon) &= \left(R(x_j + \varepsilon) + w(\varepsilon) \underline{b}(x_j + \varepsilon) \underline{b}(x_j + \varepsilon)^T \right)^{-1} \underline{b}(x_j + \varepsilon) \\ &= \left(R(x_j + \varepsilon)^{-1} - \frac{w(\varepsilon) (R(x_j + \varepsilon)^{-1} \underline{b}(x_j + \varepsilon)) (R(x_j + \varepsilon)^{-1} \underline{b}(x_j + \varepsilon))^T}{1 + w(\varepsilon) \underline{b}(x_j + \varepsilon)^T R(x_j + \varepsilon)^{-1} \underline{b}(x_j + \varepsilon)} \right) \underline{b}(x_j + \varepsilon) \\ &= \frac{R(x_j + \varepsilon)^{-1} \underline{b}(x_j + \varepsilon)}{1 + w(\varepsilon) \underline{b}(x_j + \varepsilon)^T R(x_j + \varepsilon)^{-1} \underline{b}(x_j + \varepsilon)} \longrightarrow (0, \dots, 0)^T, \quad \varepsilon \rightarrow 0. \end{aligned}$$

\square

Theorem 2.1 Zu $f \in C(K; \mathbb{R})$ sei $\varphi_i \in C^{\min\{m, l\}}(K; \mathbb{R})$ eine Kernfunktion konstruiert im Sinne von Lemma 2.1 mit $1 \leq \min\{m, l\}$. Dann gilt für jeden Multi-Index μ von der Ordnung $|\mu| = 1$

$$D^\mu \varphi_i(x_i) = \frac{1}{\sigma(x_i, x_i)} \underline{b}(x_i)^T R(x_i)^{-1} D^\mu \underline{b}(x_i) \quad (2.3)$$

und

$$D^\mu \varphi_i(x_j) = w^{(i)}(x_j) \underline{b}(x_i)^T \left(R(x_j)^{-1} - \frac{1}{\sigma(x_j, x_j)} (R(x_j)^{-1} \underline{b}(x_j)) (R(x_j)^{-1} \underline{b}(x_j))^T \right) D^\mu \underline{b}(x_j), \quad (2.4)$$

wann immer $i \neq j$ ist. Hierbei ist $w^{(i)}(x_j) = w(x_j - x_i)$ und es sind R sowie σ als Abbildungen wie in (1.43) bzw. (1.44) gegeben.

Beweis. Nach (1.62) und im Sinne von Satz 1.10 ist

$$\varphi_i(x) = w(x - x_i) \underline{b}(x_i)^T \left(BW(x) B^T \right)^{-1} \underline{b}(x), \quad x \in K,$$

so dass wir für $1 \leq \min\{m, l\}$ unter Anwendung von Lemma 1.5 schon

$$\begin{aligned} D^\mu \varphi_i(x) &= D^\mu w(x - x_i) \underline{b}(x_i)^T \left(BW(x) B^T \right)^{-1} \underline{b}(x) \\ &\quad + w(x - x_i) \underline{b}(x_i)^T D^\mu \left(BW(x) B^T \right)^{-1} \underline{b}(x) \\ &\quad + w(x - x_i) \underline{b}(x_i)^T \left(BW(x) B^T \right)^{-1} D^\mu \underline{b}(x) \\ &= D^\mu w(x - x_i) \underline{b}(x_i)^T \left(BW(x) B^T \right)^{-1} \underline{b}(x) \\ &\quad - w(x - x_i) \underline{b}(x_i)^T \left(BW(x) B^T \right)^{-1} \left(BD^\mu W(x) B^T \right) \left(BW(x) B^T \right)^{-1} \underline{b}(x) \\ &\quad + w(x - x_i) \underline{b}(x_i)^T \left(BW(x) B^T \right)^{-1} D^\mu \underline{b}(x) \end{aligned} \quad (2.5)$$

erhalten. Zunächst zeigen wir, dass für die Funktion

$$\begin{aligned} h_i(x) &:= D^\mu w(x - x_i) \underline{b}(x_i)^T \left(BW(x) B^T \right)^{-1} \underline{b}(x) \\ &\quad - w(x - x_i) \underline{b}(x_i)^T \left(BW(x) B^T \right)^{-1} \left(BD^\mu W(x) B^T \right) \left(BW(x) B^T \right)^{-1} \underline{b}(x) \end{aligned}$$

insbesondere $\lim_{\varepsilon \rightarrow 0} h_i(x_j + \varepsilon) = 0$ für alle $1 \leq j \leq n$ gilt. Im Folgenden sei $v = \frac{D^\mu w}{w}$.

Aufgrund von

$$D^\mu W(x) B^T \left(BW(x) B^T \right)^{-1} \underline{b}(x) = \begin{pmatrix} v(x - x_1) \varphi_1(x) \\ v(x - x_2) \varphi_2(x) \\ \vdots \\ v(x - x_n) \varphi_n(x) \end{pmatrix}$$

und

$$w(x - x_i) \underline{b}(x_i)^T \left(BW(x) B^T \right)^{-1} B = \left(\tilde{\varphi}_i(x, x_1), \tilde{\varphi}_i(x, x_2), \dots, \tilde{\varphi}_i(x, x_n) \right),$$

für die Funktionen

$$\tilde{\varphi}_i(x, y) = w(x - x_i) \underline{b}(x_i)^T \left(BW(x) B^T \right)^{-1} \underline{b}(y),$$

kann h_i auch in der Form

$$h_i(x) = v(x - x_i) \varphi_i(x) - \sum_{l=1}^n v(x - x_l) \varphi_l(x) \tilde{\varphi}_i(x, x_l)$$

geschrieben werden. Setzen wir hier $x = x_i + \varepsilon$, ergibt sich

$$\begin{aligned} h_i(x_i + \varepsilon) &= v(\varepsilon) \varphi_i(x_i + \varepsilon) - \sum_{l=1}^n v(x_i - x_l + \varepsilon) \varphi_l(x_i + \varepsilon) \tilde{\varphi}_i(x_i + \varepsilon, x_l) \\ &= v(\varepsilon) \varphi_i(x_i + \varepsilon) - v(\varepsilon) \varphi_i(x_i + \varepsilon) \tilde{\varphi}_i(x_i + \varepsilon, x_i) - \sum_{l \neq i} v(x_i - x_l + \varepsilon) \varphi_l(x_i + \varepsilon) \tilde{\varphi}_i(x_i + \varepsilon, x_l) \\ &= v(\varepsilon) w(\varepsilon) \underline{b}(x_i)^T \left(BW(x_i + \varepsilon) B^T \right)^{-1} \underline{b}(x_i + \varepsilon) \left(1 - w(\varepsilon) \underline{b}(x_i)^T \left(BW(x_i + \varepsilon) B^T \right)^{-1} \underline{b}(x_i) \right) \\ &\quad - \sum_{l \neq i} v(x_i - x_l + \varepsilon) \varphi_l(x_i + \varepsilon) \tilde{\varphi}_i(x_i + \varepsilon, x_l) \\ &= D^\mu w(\varepsilon) \underbrace{\underline{b}(x_i)^T \left(BW(x_i + \varepsilon) B^T \right)^{-1} \underline{b}(x_i + \varepsilon)}_{\approx \frac{1}{w(\varepsilon)}} \underbrace{\left(1 - w(\varepsilon) \underline{b}(x_i)^T \left(BW(x_i + \varepsilon) B^T \right)^{-1} \underline{b}(x_i) \right)}_{\approx \frac{\alpha}{w(\varepsilon)}} \\ &\quad - \sum_{l \neq i} v(x_i - x_l + \varepsilon) \varphi_l(x_i + \varepsilon) \tilde{\varphi}_i(x_i + \varepsilon, x_l) \\ &\approx \alpha \frac{D^\mu w(\varepsilon)}{w(\varepsilon)^2} - \sum_{l \neq i} v(x_i - x_l + \varepsilon) \varphi_l(x_i + \varepsilon) \tilde{\varphi}_i(x_i + \varepsilon, x_l) \longrightarrow 0, \quad \varepsilon \rightarrow 0, \end{aligned}$$

aufgrund von $\tilde{\varphi}_i(x_i, x_k) = 0$ für $i \neq k$ und insbesondere $\lim_{\varepsilon \rightarrow 0} \frac{D^\mu w(\varepsilon)}{w(\varepsilon)^2} = 0$ nach Definition. Hierbei ist zum einen

$$\lim_{\varepsilon \rightarrow 0} w(\varepsilon) \underline{b}(x_i)^T \left(BW(x_i + \varepsilon) B^T \right)^{-1} \underline{b}(x_i + \varepsilon) = 1,$$

das heißt $\underline{b}(x_i)^T \left(BW(x_i + \varepsilon) B^T \right)^{-1} \underline{b}(x_i + \varepsilon) \approx \frac{1}{w(\varepsilon)}$ für $\varepsilon \approx 0$ und zum anderen können wir aus der Existenz des Grenzwertes

$$\lim_{\varepsilon \rightarrow 0} w(\varepsilon) \left(1 - w(\varepsilon) \underline{b}(x_i)^T \left(BW(x_i + \varepsilon) B^T \right)^{-1} \underline{b}(x_i) \right) \quad (2.6)$$

schon $1 - w(\varepsilon) \underline{b}(x_i)^T \left(BW(x_i + \varepsilon) B^T \right)^{-1} \underline{b}(x_i) \approx \frac{\alpha}{w(\varepsilon)}$ für ein reelles α folgern.

Der Grenzwert (2.6) hingegen existiert aufgrund von

$$\begin{aligned}
 & w(\varepsilon) \left(1 - w(\varepsilon) \underline{b}(x_i)^T \left(BW(x_i + \varepsilon) B^T \right)^{-1} \underline{b}(x_i) \right) \\
 &= w(\varepsilon) - w(\varepsilon)^2 \underline{b}(x_i)^T \left(BW(x_i + \varepsilon) B^T \right)^{-1} \underline{b}(x_i) \\
 &= w(\varepsilon) - w(\varepsilon)^2 \underline{b}(x_i)^T \left(R(x_i + \varepsilon)^{-1} - \frac{w(\varepsilon) (R(x_i + \varepsilon)^{-1} \underline{b}(x_i)) (R(x_i + \varepsilon)^{-1} \underline{b}(x_i))^T}{1 + w(\varepsilon) \underline{b}(x_i)^T R(x_i + \varepsilon)^{-1} \underline{b}(x_i)} \right) \underline{b}(x_i) \\
 &= \frac{1}{1/w(\varepsilon) + \underline{b}(x_i)^T R(x_i + \varepsilon)^{-1} \underline{b}(x_i)} \longrightarrow \frac{1}{\sigma(x_i, x_i)}, \quad \varepsilon \rightarrow 0.
 \end{aligned}$$

Schließlich ist für $x = x_j + \varepsilon$ mit $i \neq j$

$$h_i(x_j + \varepsilon) = v(x_j - x_i + \varepsilon) \varphi_i(x_j + \varepsilon) - \sum_{l=1}^n v(x_j - x_l + \varepsilon) \varphi_l(x_j + \varepsilon) \tilde{\varphi}_i(x_j + \varepsilon, x_l).$$

Aufgrund von $\lim_{\varepsilon \rightarrow 0} \varphi_i(x_j + \varepsilon) = 0$ müssen wir lediglich die Gültigkeit von

$$\lim_{\varepsilon \rightarrow 0} v(\varepsilon) \varphi_j(x_j + \varepsilon) \tilde{\varphi}_i(x_j + \varepsilon, x_j) = 0$$

zeigen. Mit analogen Argumenten wie im Fall $i = j$ erhalten wir für $\varepsilon \approx 0$

$$\begin{aligned}
 & |v(\varepsilon) \varphi_j(x_j + \varepsilon) \tilde{\varphi}_i(x_j + \varepsilon, x_j)| \\
 &= |v(\varepsilon) w(\varepsilon) \underbrace{\underline{b}(x_j)^T \left(BW(x_j + \varepsilon) B^T \right)^{-1} \underline{b}(x_j)}_{\approx \frac{1}{w(\varepsilon)}} w(x_j - x_i + \varepsilon) \underline{b}(x_i)^T \left(BW(x_j + \varepsilon) B^T \right)^{-1} \underline{b}(x_j)| \\
 &\leq |v(\varepsilon) w(\varepsilon) \underbrace{\underline{b}(x_j)^T \left(BW(x_j + \varepsilon) B^T \right)^{-1} \underline{b}(x_j)}_{\approx \frac{1}{w(\varepsilon)}} \frac{w(x_j - x_i + \varepsilon)}{w(\varepsilon)} \frac{\underline{b}(x_i)^T R(x_j + \varepsilon)^{-1} \underline{b}(x_j + \varepsilon)}{\underline{b}(x_j + \varepsilon)^T R(x_j + \varepsilon)^{-1} \underline{b}(x_j + \varepsilon)}|
 \end{aligned}$$

mit Rechnungen ähnlich denen im Beweis zu Satz 2.1 und $\frac{v}{w} = \frac{D^\mu w}{w^2}$, so dass wir ebenfalls $\lim_{\varepsilon \rightarrow 0} h_i(x_j + \varepsilon) = 0$ folgern können. Aus den bewiesenen Resultaten erhalten wir

$$D^\mu \varphi_i(x_j) = \lim_{x \rightarrow x_j} \left(w(x - x_i) \underline{b}(x_i)^T \left(BW(x) B^T \right)^{-1} D^\mu \underline{b}(x) \right)$$

für beliebiges $1 \leq j \leq n$. Ist nun $i \neq j$, so liefert Lemma 1.4 unmittelbar (2.4) und für $i = j$ setzen wir in die Gleichung

$$w(x - x_i) \underline{b}(x_i)^T \left(BW(x) B^T \right)^{-1} D^\mu \underline{b}(x)$$

den Wert $x = x_i + \varepsilon$ und erhalten

$$w(\varepsilon) \underline{b}(x_i)^T \left(BW(x_i + \varepsilon) B^T \right)^{-1} D^\mu \underline{b}(x_i + \varepsilon).$$

Elementare Rechnungen liefern

$$\begin{aligned}
 & w(\varepsilon) \underline{b}(x_i)^T \left(BW(x_i + \varepsilon) B^T \right)^{-1} D^\mu \underline{b}(x_i + \varepsilon) \\
 &= w(\varepsilon) \underline{b}(x_i)^T \left(R(x_i)^{-1} - \frac{w(\varepsilon) \left(R(x_i)^{-1} \underline{b}(x_i) \right) \left(R(x_i)^{-1} \underline{b}(x_i) \right)^T}{1 + w(\varepsilon) \underline{b}(x_i)^T R(x_i)^{-1} \underline{b}(x_i)} \right) D^\mu \underline{b}(x_i + \varepsilon) \\
 &= w(\varepsilon) \underline{b}(x_i)^T \bullet \dots \\
 &\quad \bullet \frac{R(x_i)^{-1} + w(\varepsilon) \underline{b}(x_i)^T R(x_i)^{-1} \underline{b}(x_i) R(x_i)^{-1} - w(\varepsilon) \left(R(x_i)^{-1} \underline{b}(x_i) \right) \left(R(x_i)^{-1} \underline{b}(x_i) \right)^T}{1 + w(\varepsilon) \underline{b}(x_i)^T R(x_i)^{-1} \underline{b}(x_i)} \bullet \\
 &\quad \dots \bullet D^\mu \underline{b}(x_i + \varepsilon) \\
 &= \frac{1}{1/w(\varepsilon) + \sigma(x_i, x_i)} \underline{b}(x_i)^T R(x_i)^{-1} D^\mu \underline{b}(x_i + \varepsilon) \longrightarrow \frac{1}{\sigma(x_i, x_i)} \underline{b}(x_i)^T R(x_i)^{-1} D^\mu \underline{b}(x_i),
 \end{aligned}$$

mit $\varepsilon \rightarrow 0$, was den Beweis beendet. \square

2.2 Exakte kleinste Quadrate Interpolierende zur Konstruktion von Finite Differenzen

Innerhalb dieses Abschnittes gehen wir der Konstruktion von so genannten *Finite Differenzen* zur Approximation von Ableitungswerten von exakte kleinste Quadrate Interpolierenden nach. Als Konvention nehmen wir ab sofort an, dass eine äquidistante Verteilung der Interpolationsknoten vorliegt. Demnach existiert ein $\Delta x \in \mathbb{R}_{>0}$ mit

$$|x_{i+1} - x_i| = \Delta x, \quad 1 \leq i \leq n-1.$$

Zu Beginn dieses Kapitels haben wir festgestellt, dass für genügend glatte Interpolierende $\mathcal{G}f$ und Differentiationsordnungen $r \leq \min\{l, m\}$

$$\frac{d^r}{dx^r} \mathcal{G}f(x) = \sum_{i=1}^n f(x_i) \frac{d^r}{dx^r} \varphi_i(x) \quad (2.7)$$

gilt. Setzen wir hier $x = x_j$ für ein $x_j \in \{x_1, \dots, x_n\}$, lassen sich die Werte

$$\frac{d}{dx} \mathcal{G}f(x_j)$$

mit Hilfe von Theorem 2.1 aufgrund der Darstellung (2.7) bestimmen. Eine genauere Betrachtung von (2.3) sowie (2.4) legt die Vermutung nahe, dass eine Einschränkung des Trägers des unterliegenden positiven Interpolationsgewichtes auch eine entsprechend angepasste Darstellung von $\frac{d}{dx} \mathcal{G}f(x_j)$ mit sich bringt. Zur näheren Untersuchung dieser Idee bringen wir ein Beispiel.

Beispiel. Für fest gewähltes $x_j \in \{x_2, \dots, x_{n-1}\}$ und unter der Voraussetzung

$$\text{supp } w \subset \left\{ x \mid |x| \leq \frac{3}{2} \Delta x \right\}$$

gilt unter Beachtung von (2.4)

$$\begin{aligned} \frac{d}{dx} \mathcal{G}f(x_j) &= \sum_{i=j-1}^{j+1} f(x_i) \frac{d}{dx} \varphi_i(x_j) \\ &= f(x_{j-1}) \frac{d}{dx} \varphi_{j-1}(x_j) + f(x_j) \frac{d}{dx} \varphi_j(x_j) + f(x_{j+1}) \frac{d}{dx} \varphi_{j+1}(x_j). \end{aligned} \quad (2.8)$$

Mit Blick auf die Anwendbarkeit dieses Ausdrucks innerhalb des Kontextes der numerischen Approximation von Differentialgleichung ist sicherlich dem Term (2.8) gegenüber (2.7) aufgrund seiner „Lokalität“ Vorrang zu gewähren. Angenommen, der Konstruktion der Kernfunktionen in Gleichung (2.8) liegt eine quadratische monomiale Interpolationsbasis $\{e_0, e_1, e_2\}$ mit $e_k(x) = x^k$ zugrunde. Dann ist zur Berechnung der Werte $\frac{d}{dx} \varphi_i(x_j)$ für $j-1 \leq i \leq j+1$ eine Auswertung des Ausdrucks $R(x_j)^{-1}$, also die Invertierung der symmetrischen Matrix $R(x_j)$, zu gewährleisten. In unserem Fall ist diese Matrix durch

$$R(x_j) = \begin{pmatrix} w^{(j-1)} + w^{(j+1)} & x_{j-1}w^{(j-1)} + x_{j+1}w^{(j+1)} & x_{j-1}^2w^{(j-1)} + x_{j+1}^2w^{(j+1)} \\ x_{j-1}w^{(j-1)} + x_{j+1}w^{(j+1)} & x_{j-1}^2w^{(j-1)} + x_{j+1}^2w^{(j+1)} & x_{j-1}^3w^{(j-1)} + x_{j+1}^3w^{(j+1)} \\ x_{j-1}^2w^{(j-1)} + x_{j+1}^2w^{(j+1)} & x_{j-1}^3w^{(j-1)} + x_{j+1}^3w^{(j+1)} & x_{j-1}^4w^{(j-1)} + x_{j+1}^4w^{(j+1)} \end{pmatrix} \quad (2.9)$$

gegeben, wobei wir abkürzend $w^{(j-1)} \equiv w^{(j-1)}(x_j)$ und $w^{(j+1)} \equiv w^{(j+1)}(x_j)$ verwenden. Eine elementare Berechnung der Determinante von $R(x_j)$ zeigt, dass die Matrix singulär ist. Verwenden wir hingegen eine lineare monomiale Interpolationsbasis, das heißt die Menge der Funktionen $\{e_0, e_1\}$, dann ist die Matrix $R(x_j)$ durch

$$R(x_j) = \begin{pmatrix} w^{(j-1)} + w^{(j+1)} & x_{j-1}w^{(j-1)} + x_{j+1}w^{(j+1)} \\ x_{j-1}w^{(j-1)} + x_{j+1}w^{(j+1)} & x_{j-1}^2w^{(j-1)} + x_{j+1}^2w^{(j+1)} \end{pmatrix}, \quad (2.10)$$

gegeben. Im Gegensatz zu (2.9) ist (2.10) regulär.

Diese Betrachtungen legen es nahe, dass es einer genauen Analyse dieser Träger-Einschränkung des verwendeten Interpolationsgewichtes bedarf. Das nachstehende Theorem untersucht die Matrix $R(x_j)$ ausgehend von den verwendeten Interpolationsbasisfunktionen, dem positiven Interpolationsgewicht und dessen Träger-Einschränkung auf Invertierbarkeit. Der Leser ist frei in seiner Wahl, das Theorem als univariates oder multivariates Resultat zu interpretieren; der Beweis verläuft in beiden Fällen analog.

Theorem 2.2 Für $1 \leq i \leq n$ sei $m^{(i)} \in C^l(K; \mathbb{R})$ eine Funktion mit $\text{supp } m^{(i)} = D_i$ derart, dass $x_i \in \text{int } D_i$ und $m^{(i)}(x) > 0$ für alle $x \in \text{int } D_i$ ist. Darüber hinaus sei

$$w^{(i)}(x) = m^{(i)}(x) \cdot w(x - x_i), \quad x \in K, \quad 1 \leq i \leq n,$$

für ein positives Interpolationsgewicht $w \in C^l(\mathbb{R}^d \setminus \{0\}; \mathbb{R}_{>0})$. Die Matrix $R(x_j) \in \mathbb{R}^{k \times k}$ sei durch

$$R(x_j) = \begin{pmatrix} \sum_{\substack{i=1 \\ i \neq j}}^n b^{(1)}(x_i) w^{(i)}(x_j) b^{(1)}(x_i) & \dots & \sum_{\substack{i=1 \\ i \neq j}}^n b^{(1)}(x_i) w^{(i)}(x_j) b^{(k)}(x_i) \\ \vdots & \ddots & \vdots \\ \sum_{\substack{i=1 \\ i \neq j}}^n b^{(k)}(x_i) w^{(i)}(x_j) b^{(1)}(x_i) & \dots & \sum_{\substack{i=1 \\ i \neq j}}^n b^{(k)}(x_i) w^{(i)}(x_j) b^{(k)}(x_i) \end{pmatrix}$$

definiert, wobei durch $\{b^{(1)}, \dots, b^{(k)}\}$ eine Interpolationsbasis gegeben ist.

Angenommen, es existiert eine Zerlegung $\{i_1, \dots, i_r\} \cup \{i_{r+1}, \dots, i_n\}$ der Indexmenge $\{1, \dots, n\}$, so dass

$$x_j \in \bigcap_{k=1}^r D_{i_k}, \quad x_j \notin \bigcup_{k=r+1}^n D_{i_k} \quad (2.11)$$

erfüllt ist. Dann ist die Matrix $R(x_j)$ regulär für $r > k$ und singulär für $r \leq k$.

Anhand eines Beispiels gehen wir im Anschluss an den Beweis von Theorem 2.2 näher auf die so genannte *Überdeckungs-Eigenschaft* (2.11) ein.

Beweis von Theorem 2.2. Für $x_j \in \{x_1, \dots, x_n\}$ existiert eine Menge von Indizes

$$J_r := \{i_1, \dots, i_r\} \setminus \{j\}$$

mit $x_j \in \bigcap_{k \in J_r} D_{i_k}$ und $m^{(i)}(x_j) = 0$ für alle $i \in \{1, \dots, n\} \setminus (J_r \cup \{j\})$. Hieraus können wir

$$\begin{aligned} R(x_j) &= \begin{pmatrix} \sum_{i \in J_r} b^{(1)}(x_i) w^{(i)}(x_j) b^{(1)}(x_i) & \dots & \sum_{i \in J_r} b^{(1)}(x_i) w^{(i)}(x_j) b^{(k)}(x_i) \\ \vdots & \ddots & \vdots \\ \sum_{i \in J_r} b^{(k)}(x_i) w^{(i)}(x_j) b^{(1)}(x_i) & \dots & \sum_{i \in J_r} b^{(k)}(x_i) w^{(i)}(x_j) b^{(k)}(x_i) \end{pmatrix} \\ &= \sum_{i \in J_r} w^{(i)}(x_j) \underline{b}(x_i) \underline{b}(x_i)^T \end{aligned}$$

schließen. Die dyadischen Produkte $\underline{b}(x_i) \underline{b}(x_i)^T$ sind Matrizen vom Rang Eins, das heißt es ist $\text{Rang}(\underline{b}(x_i) \underline{b}(x_i)^T) = 1$. Wir untersuchen jetzt entsprechend des Zusammenhanges von r und k die Invertierbarkeit von $R(x_j)$.

Fall I ($r \leq k$): Da $R(x_j)$ die Summe von $r - 1 < k$ Matrizen vom Rang Eins ist, gilt

$$\text{Rang} R(x_j) \leq (r - 1) \cdot 1 < k,$$

das heißt $R(x_j)$ ist singulär.

Fall II ($r = k + 1$): Nehmen wir an, $R(x_j)$ wäre singulär. Dann existieren $\alpha_1, \dots, \alpha_k \in \mathbb{R}$ mit

$$\begin{aligned} \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} &= \sum_{l=1}^k \alpha_l \sum_{i \in J_r} w^{(i)}(x_j) b^{(l)}(x_i) \begin{pmatrix} b^{(1)}(x_i) \\ \vdots \\ b^{(k)}(x_i) \end{pmatrix} \\ &= \sum_{i \in J_r} \left(w^{(i)}(x_j) \sum_{l=1}^k \alpha_l b^{(l)}(x_i) \right) \begin{pmatrix} b^{(1)}(x_i) \\ \vdots \\ b^{(k)}(x_i) \end{pmatrix}. \end{aligned}$$

Dies steht jedoch im Widerspruch zur Annahme, dass die Interpolationsbasisfunktionen linear unabhängig auf jeder Teilmenge $\{x_{i_1}, \dots, x_{i_k}\} \subset \{x_1, \dots, x_n\}$ sind.

Fall III ($r > k + 1$): Die Regularität von $R(x_j)$ liefert unmittelbar ein Blick auf den Fall II, die Unabhängigkeit der Interpolationsbasisfunktionen und eine Anwendung von Standardregeln zur Rangtransformation unter elementaren Matrix-Umformungen. \square

Bezugnehmend auf unser Beispiel (2.8) ist an dieser Stelle zu überlegen, von welcher Bauart die Funktionen $m^{(i)}$ in Theorem 2.2 zu wählen sind, um das aufgezeigte Resultat rigoros zu erhalten. Definieren wir hierfür die Mengen U_i und V_i durch

$$U_i := \left\{ x \in \mathbb{R} \left| |x - x_i| \leq \frac{3}{2} \Delta x \right. \right\}$$

beziehungsweise

$$V_i := \left\{ x \in \mathbb{R} \left| |x - x_i| < \frac{7}{4} \Delta x \right. \right\},$$

so existiert ein $m^{(i)} \in C^l(\mathbb{R}; \mathbb{R})$ mit $m^{(i)}(x) = 1$ für $x \in U_i$ und $m^{(i)}(x) = 0$ für $x \in \mathbb{R} \setminus V_i$, das heißt es ist

$$\frac{d}{dx} \varphi_i(x_j) = 0, \quad i = 1, \dots, j-2, j+2, \dots, n,$$

und wir erhalten (2.8). Dabei ist einzusehen, dass die tatsächlichen Definitionen der Mengen U_i und V_i unwesentlich sind. Innerhalb dieses Kontextes sind vielmehr die Eigenschaften $x_{i-1}, x_i, x_{i+1} \in U_i$ und $\{x_1, \dots, x_{i-2}, x_{i+2}, \dots, x_n\} \subset \mathbb{R} \setminus V_i$ essentiell.

Möchten wir allgemeiner eine Darstellung der Form

$$\frac{d}{dx} \mathcal{G}f(x_j) = \sum_{i=j-s}^{j+s} f(x_i) \frac{d}{dx} \varphi_i(x_j) \quad (2.12)$$

erreichen, können die zur Konstruktion der Funktionen $m^{(i)}$ heranzuziehenden Mengen U_i und V_i beispielsweise als

$$U_i := \left\{ x \in \mathbb{R} \left| |x - x_i| \leq \frac{2s+1}{2} \Delta x \right. \right\}$$

und

$$V_i := \left\{ x \in \mathbb{R} \left| |x - x_i| < \frac{4s+3}{4} \Delta x \right. \right\}$$

gewählt werden. Die Funktionen $m^{(i)}$ haben hierbei wieder $m^{(i)}(x) = \begin{cases} 1 & : x \in U_i \\ 0 & : x \in \mathbb{R} \setminus V_i \end{cases}$ zu erfüllen.

Definition 2.1 Der Ausdruck

$$\sum_{i=j-s}^{j+s} f(x_i) \frac{d}{dx} \varphi_i(x_j)$$

wird *symmetrische* $(2s + 1)$ -Punkte Finite Differenz genannt. Für (2.12) wird auch

$$\frac{d}{dx} \mathcal{G}f(x_j) \sim \boxed{\varphi'_{j-s}(x_j)} \bullet \cdots \bullet \boxed{\varphi'_j(x_j)} \bullet \cdots \bullet \boxed{\varphi'_{j+s}(x_j)},$$

in Übereinstimmung mit Standardnotationen, geschrieben.

Der Begriff „symmetrisch“ ist hierbei nicht in einer Weise zu interpretieren, dass die Finite Differenz von der Form

$$\frac{d}{dx} \mathcal{G}f(x_l) \sim \left[\boxed{\lambda_0} \bullet \cdots \bullet \boxed{\lambda_1} \bullet \cdots \bullet \boxed{\lambda_0} \right],$$

ist. Vielmehr verwenden wir den Begriff „symmetrisch“, um die Struktur der Finite Differenz nach Definition 2.1 wiederzugeben.

An dieser Stelle kommt die Frage auf, ob die Approximation des Wertes $\frac{d}{dx} \mathcal{G}f(x_j)$ durch

$$\sum_{i=j-s}^{j+s} f(x_i) \frac{d}{dx} \varphi_i(x_j) \quad (2.13)$$

wiederum einen Rückschluss auf eine etwaige Approximation von $\frac{d}{dx} f(x_j)$ durch (2.13) zulässt, falls die Funktion f entsprechend differenzierbar ist. Auf diesen Aspekt, der *konservative* bzw. *simultane Approximation* genannt wird, geht Kapitel 4 näher ein.

2.3 Inverse Euklidische Normen als positive Interpolationsgewichte

Bis zu diesem Zeitpunkt konnten wir innerhalb der vorliegenden Arbeit im Wesentlichen auf die genaue Angabe von positiven Interpolationsgewichten verzichten. Vielmehr erschien die explizite Struktur der Gewichte sogar unwesentlich, so lange wir deren Eigenschaften nach Definition 1.4 nutzen konnten. Die folgende Definition stellt eine Klasse von positiven Interpolationsgewichten bereit, mit denen wir im weiteren Verlauf arbeiten werden.

Definition 2.2 Sei $w \in C^l(\mathbb{R}^d \setminus \{0\}; \mathbb{R}_{>0})$ ein positives Interpolationsgewicht. Ferner seien die Zahlen $\alpha, \beta \in \mathbb{N}$ und $\mathfrak{a}, \mathfrak{b} \in \mathbb{R}_{>0}$ beliebig aber fest gewählt. Dann heißt das Gewicht

- (i) *symmetrisch* und *ungewichtet Norm-invers*, falls

$$w \equiv w[\alpha](x) = \|x\|_2^{-\alpha},$$

und speziell im Fall $d = 1$

- (ii) *unsymmetrisch* und *ungewichtet Norm-invers*, falls

$$w \equiv w[\alpha|\beta](x) = \begin{cases} |x|^{-\alpha} & : x < 0 \\ |x|^{-\beta} & : x > 0 \end{cases},$$

(iii) *unsymmetrisch* und *gewichtet Norm-invers*, falls

$$w \equiv w[\alpha, \mathfrak{a}|\beta, \mathfrak{b}](x) = \begin{cases} \mathfrak{a} \cdot |x|^{-\alpha} & : x < 0 \\ \mathfrak{b} \cdot |x|^{-\beta} & : x > 0 \end{cases},$$

gilt.

Die unsymmetrischen Norm-inversen positiven Interpolationsgewichte spielen bei der Konstruktion von Finite Differenzen zur Approximation von Ableitungswerten im Kontext der numerischen Behandlung von nichtlinearen Advektionsgleichungen eine wesentliche Rolle. Wir kommen in Kapitel 5 hierauf zurück.

2.4 Die Konstruktion und Analyse symmetrischer 3-Punkte Finite Differenzen

Im Folgenden bringen wir einige Resultate von Finite Differenzen-Approximationen im Sinne von Abschnitt 2.2, deren Konstruktion auf exakte kleinste Quadrate Interpolierenden mit zugrundeliegender linearer monomialer Interpolationsbasis der Form $\{e_0, e_1\}$ mit $e_k(x) = x^k$ und Norm-inversen positiven Interpolationsgewichten nach Definition 2.2 basiert. Wir beschränken uns innerhalb dieses Abschnitts auf 3-Punkte Finite Differenzen und verweisen auf Kapitel 5, welches weitere Beispiele liefert.

Wir nehmen zudem äquidistant verteilte Interpolationsknoten $\{x_1, \dots, x_n\} \subset \mathbb{R}$ an, das heißt es existiert ein Δx mit $|x_{i+1} - x_i| = \Delta x$ für $1 \leq i \leq n-1$. Ferner sei $x_j \in \{x_2, \dots, x_{n-1}\}$ fest gewählt. Weiter betrachten wir unsymmetrische und ungewichtete Norm-inverse positive Interpolationsgewichte, also Gewichte von der Form

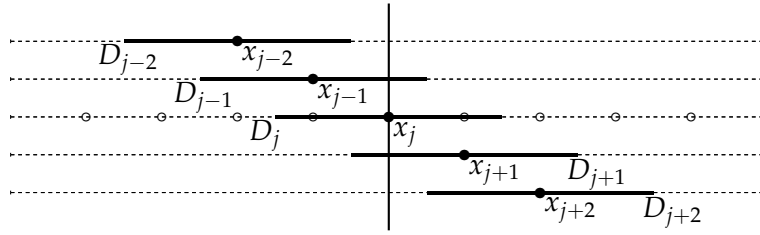
$$w \equiv w[\alpha|\beta](x) = \begin{cases} |x|^{-\alpha} & : |x| < 0 \\ |x|^{-\beta} & : |x| > 0 \end{cases}, \quad \alpha, \beta \in \mathbb{N}.$$

Unser Ziel ist die Konstruktion einer symmetrischen 3-Punkte Finite Differenz zur Approximation des Wertes $\frac{d}{dx} \mathcal{G}f(x_j)$. Zunächst ist zu überprüfen, ob die Konstruktion mit zugrundeliegender linearer Interpolationsbasis durchführbar ist: Formal erhalten wir die angestrebte 3-Punkte Finite Differenz, indem die unterliegende exakte kleinste Quadrate Interpolierende mit passend gewählten Träger-eingeschränkten Translaten $w^{(i)} = m^{(i)} \cdot w(\cdot - x_i)$, $1 \leq i \leq n$, des positiven Interpolationsgewichtes w konstruiert wird. Wir verzichten an dieser Stelle auf eine explizite Angabe der Funktionen $m^{(i)}$ und verweisen auf die Erörterungen im Anschluss an Theorem 2.2.

Ein Blick auf Abbildung 2.1 sowie Theorem 2.2 verrät, dass die Konstruktion der Finite Differenz möglich ist, denn je nach Beschaffenheit der Träger D_i existiert offensichtlich ein $r > 2 = k$, so dass die gewünschte Überdeckungseigenschaft nach Theorem 2.2 erfüllt ist.

Für die Berechnung der symmetrischen 3-Punkte Finite Differenz notieren wir $\underline{b}(x_j) = (1, x_j)^T$ und $\frac{d}{dx} \underline{b}(x_j) = (0, 1)^T$. Weiter verwenden wir die Abkürzung $w^{(i)}(x) = w(x - x_i)$ für $1 \leq i \leq n$. Ein Blick auf die Identitäten

$$\begin{aligned} w^{(j-1)}(x_j) &= |x_j - x_{j-1}|^{-\beta} = \Delta x^{-\beta}, \\ w^{(j+1)}(x_j) &= |x_j - x_{j+1}|^{-\alpha} = \Delta x^{-\alpha}, \end{aligned}$$

Abbildung 2.1: Test der Überdeckungseigenschaft für $k = 2$ und 3-Punkte Finite Differenzen.

liefert

$$\begin{aligned}
 R(x_j) &= \begin{pmatrix} w^{(j-1)}(x_j) + w^{(j+1)}(x_j) & x_{j-1}w^{(j-1)}(x_j) + x_{j+1}w^{(j+1)}(x_j) \\ x_{j-1}w^{(j-1)}(x_j) + x_{j+1}w^{(j+1)}(x_j) & x_{j-1}^2w^{(j-1)}(x_j) + x_{j+1}^2w^{(j+1)}(x_j) \end{pmatrix} \\
 &= \begin{pmatrix} \Delta x^{-\beta} + \Delta x^{-\alpha} & x_{j-1}\Delta x^{-\beta} + x_{j+1}\Delta x^{-\alpha} \\ x_{j-1}\Delta x^{-\beta} + x_{j+1}\Delta x^{-\alpha} & x_{j-1}^2\Delta x^{-\beta} + x_{j+1}^2\Delta x^{-\alpha} \end{pmatrix} = \frac{1}{\Delta x^{\alpha+\beta}} M,
 \end{aligned}$$

mit der Matrix

$$M = \begin{pmatrix} \Delta x^{\alpha} + \Delta x^{\beta} & x_j(\Delta x^{\alpha} + \Delta x^{\beta}) - \Delta x^{\alpha+1} + \Delta x^{\beta+1} \\ x_j(\Delta x^{\alpha} + \Delta x^{\beta}) - \Delta x^{\alpha+1} + \Delta x^{\beta+1} & x_j^2(\Delta x^{\alpha} + \Delta x^{\beta}) + 2x_j(-\Delta x^{\alpha+1} + \Delta x^{\beta+1}) + \Delta x^{\alpha+2} + \Delta x^{\beta+2} \end{pmatrix}.$$

Die Inverse der Matrix $R(x_j)$ kann mit Hilfe der Adjunkten-Regel berechnet werden.

Wir erhalten

$$\begin{aligned}
 R(x_j)^{-1} &= \frac{\Delta x^{\alpha+\beta}}{\det M} \bullet \dots \\
 \dots \bullet &\begin{pmatrix} x_j^2(\Delta x^{\alpha} + \Delta x^{\beta}) + 2x_j(-\Delta x^{\alpha+1} + \Delta x^{\beta+1}) + \Delta x^{\alpha+2} + \Delta x^{\beta+2} & -x_j(\Delta x^{\alpha} + \Delta x^{\beta}) + \Delta x^{\alpha+1} - \Delta x^{\beta+1} \\ -x_j(\Delta x^{\alpha} + \Delta x^{\beta}) + \Delta x^{\alpha+1} - \Delta x^{\beta+1} & \Delta x^{\beta} + \Delta x^{\alpha} \end{pmatrix},
 \end{aligned}$$

mit $\det M = 4\Delta x^{\alpha+\beta+2}$. Es folgt

$$\sigma(x_j, x_j) = \underline{b}(x_j)^T R(x_j)^{-1} \underline{b}(x_j) = \frac{\Delta x^{\alpha+\beta+2}}{\det M} (\Delta x^{\alpha} + \Delta x^{\beta}),$$

womit durch Theorem 2.1 sofort

$$\frac{d}{dx} \varphi_j(x_j) = \frac{1}{\sigma(x_j, x_j)} \underline{b}(x_j)^T R(x_j)^{-1} \frac{d}{dx} \underline{b}(x_j) = \frac{1}{\Delta x} \frac{\Delta x^{\alpha} - \Delta x^{\beta}}{\Delta x^{\alpha} + \Delta x^{\beta}}$$

resultiert. Weiter ist für $i = j \pm 1$ nach Theorem 2.1

$$\frac{d}{dx} \varphi_i(x_j) = w^{(i)}(x_j) \underline{b}(x_i)^T \left(R(x_j)^{-1} - \frac{1}{\sigma(x_j, x_j)} (R(x_j)^{-1} \underline{b}(x_j)) (R(x_j)^{-1} \underline{b}(x_j))^T \right) \frac{d}{dx} \underline{b}(x_j).$$

Für $i = j + 1$ gilt $w^{(i)}(x_j) = w^{(j+1)}(x_j) = \Delta x^{-\alpha}$ und $\underline{b}(x_{j+1}) = (1, x_j + \Delta x)^T$. Elementare, aber umfangreichere Rechnungen liefern

$$R(x_j)^{-1} \underline{b}(x_j) = \frac{\Delta x^{\alpha+\beta}}{\det M} \begin{pmatrix} x_j \Delta x (\Delta x^\beta - \Delta x^\alpha) + \Delta x^2 (\Delta x^\alpha + \Delta x^\beta) \\ \Delta x (\Delta x^\alpha - \Delta x^\beta) \end{pmatrix},$$

$$R(x_j)^{-1} - \frac{1}{\sigma(x_j, x_j)} (R(x_j)^{-1} \underline{b}(x_j)) (R(x_j)^{-1} \underline{b}(x_j))^T =$$

$$\begin{pmatrix} * & \frac{(-x_j(\Delta x^\alpha + \Delta x^\beta) + \Delta x^{\alpha+1} - \Delta x^{\beta+1}) \frac{\Delta x^{\alpha+\beta}}{\det M}}{-\frac{1}{\sigma} \left(\frac{\Delta x^{\alpha+\beta}}{\det M} \right)^2 (-x_j \Delta x^2 (\Delta x^\alpha - \Delta x^\beta)^2 + \Delta x^3 (\Delta x^{2\alpha} - \Delta x^{2\beta}))} \\ * & \frac{(\Delta x^\alpha + \Delta x^\beta) \frac{\Delta x^{\alpha+\beta}}{\det M}}{-\frac{1}{\sigma} \left(\frac{\Delta x^{\alpha+\beta}}{\det M} \right)^2 \Delta x^2 (\Delta x^\alpha - \Delta x^\beta)^2} \end{pmatrix}$$

sowie

$$\frac{d}{dx} \varphi_{j+1}(x_j) = \frac{1}{2\Delta x} \left(1 - \frac{\Delta x^\alpha - \Delta x^\beta}{\Delta x^\alpha + \Delta x^\beta} \right).$$

Hierbei sind die mit ‘*’ gekennzeichneten Matrix-Einträge bei den auftretenden Rechnungen nicht von Interesse und wir geben diese daher nicht an. Eine analoge Rechnung für $i = j - 1$ mit $w^{(i)}(x_j) = w^{(j-1)}(x_j) = \Delta x^{-\beta}$ und $\underline{b}(x_{j-1}) = (1, x_j - \Delta x)^T$ zeigt

$$\frac{d}{dx} \varphi_{j-1}(x_j) = -\frac{1}{2\Delta x} \left(1 + \frac{\Delta x^\alpha - \Delta x^\beta}{\Delta x^\alpha + \Delta x^\beta} \right).$$

Mit Hilfe der Notation aus Definition 2.1 können wir dieses Ergebnis als

$$\frac{d}{dx} \mathcal{G}f(x_j) \sim \boxed{-\frac{1}{2\Delta x} \left(1 + \frac{\Delta x^\alpha - \Delta x^\beta}{\Delta x^\alpha + \Delta x^\beta} \right)} \bullet \bullet \boxed{\frac{1}{\Delta x} \frac{\Delta x^\alpha - \Delta x^\beta}{\Delta x^\alpha + \Delta x^\beta}} \bullet \bullet \boxed{\frac{1}{2\Delta x} \left(1 - \frac{\Delta x^\alpha - \Delta x^\beta}{\Delta x^\alpha + \Delta x^\beta} \right)} \quad (2.14)$$

ausdrücken.

Ein genauerer Blick auf die Approximation (2.14) lässt zudem die Struktur

$$\begin{aligned} \frac{d}{dx} \mathcal{G}f(x_j) = & \frac{-1}{2\Delta x} \frac{\Delta x^\alpha - \Delta x^\beta}{\Delta x^\alpha + \Delta x^\beta} \cdot f_{j-1} + \frac{1}{2\Delta x} \frac{\Delta x^\alpha - \Delta x^\beta}{\Delta x^\alpha + \Delta x^\beta} \cdot f_j + 0 \cdot f_{j+1} \\ & \frac{-1}{2\Delta x} \cdot f_{j-1} + 0 \cdot f_j + \frac{1}{2\Delta x} \cdot f_{j+1} \\ & + 0 \cdot f_{j-1} + \frac{1}{2\Delta x} \frac{\Delta x^\alpha - \Delta x^\beta}{\Delta x^\alpha + \Delta x^\beta} \cdot f_j + \frac{-1}{2\Delta x} \frac{\Delta x^\alpha - \Delta x^\beta}{\Delta x^\alpha + \Delta x^\beta} \cdot f_{j+1} \end{aligned}$$

mit $f_k \equiv f(x_k)$ für $j - 1 \leq k \leq j + 1$ erkennen; diese ist aus einer Rückwärts-, zentralen und Vorwärts-Differenz zusammengesetzt. Durch Variation der Werte α und β kann der Anteil jeder einzelnen Differenz zum Gesamtausdruck mehr oder weniger stark gewichtet werden.

Schauen wir uns diese Gewichtung im Einzelfall etwas näher an: Für $\alpha = \beta$ erhalten wir die klassische zentrale Differenz

$$\frac{d}{dx} \mathcal{G}f(x_j) = \frac{-f_{j-1} + f_{j+1}}{2\Delta x}.$$

Setzen wir α beliebig aber fest und betrachten den Grenzfall $\beta \rightarrow \infty$, so resultiert

$$\frac{d}{dx} \mathcal{G}f(x_j) = \frac{-f_{j-1} + f_j}{\Delta x},$$

eine klassische Rückwärtsdifferenz. Wählen wir schlussendlich β beliebig aber fest und befinden uns im Grenzfall $\alpha \rightarrow \infty$, können wir die Identität

$$\frac{d}{dx} \mathcal{G}f(x_j) = \frac{-f_j + f_{j+1}}{\Delta x},$$

das heißt eine klassische Vorwärtsdifferenz als Finite Differenz, feststellen.

Nehmen wir a priori an, dass durch $\frac{d}{dx} \mathcal{G}f(x_j)$ auch eine Approximation für den Wert $\frac{d}{dx} f(x_j)$ gegeben ist, sofern dieser Wert existiert und die Funktion f hinreichend differenzierbar ist. Dann kann nach der *Konsistenz* der Approximation $\frac{d}{dx} \mathcal{G}f(x_j)$ bezüglich des Ableitungswertes $\frac{d}{dx} f(x_j)$ gefragt werden. Hierbei verwenden wir die Begriffe *Konsistenz* und *Konsistenzordnung* im Sinne von Richtmyer und Morton [23].

Für das obige Beispiel wird die Konsistenzfrage durch folgende Rechnung beantwortet: Eine Taylorentwicklung der Form

$$f_{j\pm 1} = f_j \pm \Delta x \frac{d}{dx} f_j + \frac{\Delta x^2}{2!} \frac{d^2}{dx^2} f_j \pm \frac{\Delta x^3}{3!} \frac{d^3}{dx^3} f_j + \dots$$

liefert

$$\begin{aligned} \underbrace{\frac{d}{dx} \mathcal{G}f(x_j)}_{\approx \frac{d}{dx} f(x_j)} - \frac{1}{\Delta x} \left(-\frac{1}{2} \left(1 + \frac{\Delta x^\alpha - \Delta x^\beta}{\Delta x^\alpha + \Delta x^\beta} \right) f_{j-1} + \frac{\Delta x^\alpha - \Delta x^\beta}{\Delta x^\alpha + \Delta x^\beta} f_j + \frac{1}{2} \left(1 - \frac{\Delta x^\alpha - \Delta x^\beta}{\Delta x^\alpha + \Delta x^\beta} \right) f_{j+1} \right) \\ = -\frac{\Delta x^2}{3!} \frac{d^3}{dx^3} f_j - \frac{\Delta x^4}{5!} \frac{d^5}{dx^5} f_j - \frac{\Delta x^6}{7!} \frac{d^7}{dx^7} f_j \\ + \frac{\Delta x}{2!} \frac{\Delta x^\alpha - \Delta x^\beta}{\Delta x^\alpha + \Delta x^\beta} \frac{d^2}{dx^2} f_j + \frac{\Delta x^3}{4!} \frac{\Delta x^\alpha - \Delta x^\beta}{\Delta x^\alpha + \Delta x^\beta} \frac{d^4}{dx^4} f_j + \dots = \begin{cases} \mathcal{O}(\Delta x) & : \alpha \neq \beta \\ \mathcal{O}(\Delta x^2) & : \alpha = \beta \end{cases}. \end{aligned}$$

Im Fall $\alpha = \beta$ ist $\frac{d}{dx} \mathcal{G}f(x_j)$ daher eine Approximation konsistent von zweiter Ordnung zu $\frac{d}{dx} f(x_j)$, da die Terme

$$\frac{\Delta x^\alpha - \Delta x^\beta}{\Delta x^\alpha + \Delta x^\beta}$$

wegfallen. Für $\alpha \neq \beta$ ist die Approximation hingegen konsistent von erster Ordnung.

Arbeiten wir mit gewichteten und unsymmetrischen Norm-inversen positiven Interpolationsgewichten, liefern analoge Rechnungen im Sinne von Theorem 2.1

$$\frac{d}{dx} \mathcal{G}f(x_j) \sim \left[\frac{-1}{2\Delta x} \left(1 + \frac{b\Delta x^\alpha - a\Delta x^\beta}{b\Delta x^\alpha + a\Delta x^\beta} \right) \right] \bullet \bullet \left[\frac{1}{\Delta x} \frac{b\Delta x^\alpha - a\Delta x^\beta}{b\Delta x^\alpha + a\Delta x^\beta} \right] \bullet \bullet \left[\frac{1}{2\Delta x} \left(1 - \frac{b\Delta x^\alpha - a\Delta x^\beta}{b\Delta x^\alpha + a\Delta x^\beta} \right) \right]. \quad (2.15)$$

Die Konsistenzanalyse für die 3-Punkte Finite Differenz basiert dann auf

$$\begin{aligned} & \underbrace{\frac{d}{dx} \mathcal{G}f(x_j)}_{\approx \frac{d}{dx} f(x_j)} \\ & - \frac{1}{\Delta x} \left(-\frac{1}{2} \left(1 + \frac{b\Delta x^\alpha - a\Delta x^\beta}{b\Delta x^\alpha + a\Delta x^\beta} \right) f_{j-1} + \frac{b\Delta x^\alpha - a\Delta x^\beta}{b\Delta x^\alpha + a\Delta x^\beta} f_j + \frac{1}{2} \left(1 - \frac{b\Delta x^\alpha - a\Delta x^\beta}{b\Delta x^\alpha + a\Delta x^\beta} \right) f_{j+1} \right) \\ & = -\frac{\Delta x^2}{3!} \frac{d^3}{dx^3} f_j - \frac{\Delta x^4}{5!} \frac{d^5}{dx^5} f_j - \frac{\Delta x^6}{7!} \frac{d^7}{dx^7} f_j + \dots \\ & \quad + \frac{\Delta x}{2!} \frac{b\Delta x^\alpha - a\Delta x^\beta}{b\Delta x^\alpha + a\Delta x^\beta} \frac{d^2}{dx^2} f_j + \frac{\Delta x^3}{4!} \frac{b\Delta x^\alpha - a\Delta x^\beta}{b\Delta x^\alpha + a\Delta x^\beta} \frac{d^4}{dx^4} f_j + \frac{\Delta x^5}{6!} \frac{b\Delta x^\alpha - a\Delta x^\beta}{b\Delta x^\alpha + a\Delta x^\beta} \frac{d^6}{dx^6} f_j + \dots \\ & = \begin{cases} \mathcal{O}(\Delta x) & : b\Delta x^\alpha \neq a\Delta x^\beta \\ \mathcal{O}(\Delta x^2) & : b\Delta x^\alpha = a\Delta x^\beta \end{cases} . \end{aligned} \quad (2.16)$$


Wir erkennen, dass hierdurch ähnliche Konsistenzordnungen wie der im Fall der Verwendung von ungewichteten und unsymmetrischen Norm-inversen positiven Interpolationsgewichten gewonnen werden.

An dieser Stelle verzichten wir auf die Angabe und Analyse weiterer Beispiele und verweisen hierfür auf Kapitel 5.

Kapitel 3

Konvergenzresultate für exakte kleinste Quadrate Interpolierende vom Korovkin-Typ

3.1 Konvergenz der univariaten Shepard-Interpolation	46
3.2 Über die Namensgebung des Konvergenzsatzes von Bohman und Korovkin	53
3.3 Konvergenz von Folgen fast positiver linearer Operatoren auf $C(K; \mathbb{R})$	54
3.4 Das Konvergenzproblem für die exakte kleinste Quadrate Interpolation	59
3.5 Konvergenz bei konstanter monomialer Interpolationsbasis	61
3.6 Eine detailliertere Untersuchung der Kernfunktionen	64
3.7 Konvergenz bei quadratischer monomialer Interpolationsbasis	69
3.8 Konvergenz bei linearer monomialer Interpolationsbasis	73

ätze vom Korovkin-Typ¹ bilden in der klassischen Approximationstheorie einen Zweig mächtiger Werkzeuge im Kontext von Konvergenzaussagen für Folgen von positiven und linearen Operatoren. Dieses Kapitel ist dem Konvergenzproblem für die exakte kleinste Quadrate Interpolation gewidmet. Wir untersuchen für monomiale Interpolationsbasen, inwiefern die dazugehörige exakte kleinste Quadrate Interpolierende einer stetigen Funktion gegen diese im Sinne einer wachsenden Anzahl der zugrundeliegenden Interpolationsknoten konvergiert. Es stellt sich heraus, dass die korrespondierenden Approximationsoperatoren im Allgemeinen nicht positiv sind, so dass wir die in der Literatur vorhandenen Konvergenzresultate für unseren Anwendungskontext auf so genannte Folgen von fast positiven Operatoren verallgemeinern müssen.

Wann immer wir innerhalb dieses Kapitels von einem positiven Interpolationsgewicht sprechen, ist ein symmetrisches und ungewichtetes Norm-inverses positives Interpolationsgewicht entsprechend der Definition 2.2 gemeint, sofern nichts anderes angegeben wird.

¹Nach Pavel Petrovich Korovkin, 1913-1985.

3.1 Konvergenz der univariaten Shepard-Interpolation

Innerhalb dieses Abschnitts untersuchen wir die univariate Shepard-Interpolation auf Konvergenz. Die aufgezeigten Methoden und Ideen liefern die Grundlage aller weiteren Resultate dieses Kapitels.

3.1.i Der Konvergenzsatz von Bohman und Korovkin

Wir bringen zunächst eine klassische Version des Konvergenzsatzes von Bohman und Korovkin und folgern aus diesem den Weierstraßschen Approximationssatz. Der Satz von Bohman und Korovkin gibt Auskunft über Bilder stetiger Funktionen unter positiven und linearen Operatoren. Hierzu führen wir zunächst den Begriff eines *positiven Operators* ein.

Definition 3.1 Seien $(X; \leq)$ sowie $(Y; \leq)$ \mathbb{K} -Vektorräume mit Ordnungsrelation \leq und es gelte die Inklusion $Y \subseteq X$. Ein Operator $P : X \rightarrow Y$ heißt *positiv*, falls

$$0 \leq f \implies 0 \leq Pf$$

für alle $f \in X$ gilt.

Ist der positive Operator $P : X \rightarrow Y$ ebenfalls mit der Eigenschaft der Linearität ausgestattet, lässt sich für diesen insbesondere Monotonie zeigen. Vielmehr sind für lineare Operatoren $L : X \rightarrow Y$ Positivität und Monotonie äquivalent: Ist L monoton, das heißt $g \leq h \implies Lg \leq Lh$, folgt die Positivität durch Setzen von $g = 0_X$. Anders herum folgt direkt aus der Positivität, das heißt $0_X \leq h - g \implies 0_Y \leq L(h - g)$, mit Hilfe der Linearität von L die Monotonie.

Für Folgen $(P_n)_{n \in \mathbb{N}}$ von linearen und positiven Operatoren $P_n : X \rightarrow Y$ existieren in der Approximationstheorie mächtige Konvergenzresultate, die von folgender Struktur sind:

Vorausgesetzt $\lim_{n \rightarrow \infty} P_n h_i$ existiert in Y für gewisse $h_1, \dots, h_s \in X$, dann existiert der Grenzwert $\lim_{n \rightarrow \infty} P_n h$ in Y für alle $h \in X$.

Auf die Frage wie die Grenzwertoperatoren im Einzelfall aussehen, gehen wir im weiteren Verlauf ein. Bei der obigen Aussage spricht man von einem Konvergenzresultat vom *Korovkin-* bzw. *Bohman-Korovkin-Typ*. Diese Bezeichnung basiert auf Ergebnissen, die Bohman [5] und Korovkin [14, 15] unabhängig voneinander, aber in ähnlicher Weise bewiesen haben. Für eine historische Einordnung dieser Ergebnisse verweisen wir auf Abschnitt 3.2.

Theorem 3.1 ist eine erste Version dieses Resultates, die sich in ihrer Darstellung an der Monographie [15] orientiert.

Theorem 3.1 (Bohman, Korovkin) Sei $(P_n)_{n \geq n_0}$ für festes $n_0 \in \mathbb{N}$ eine Folge von positiven und linearen Operatoren, die den Raum $C([a, b]; \mathbb{R})$ in sich selbst abbilden. Vorausgesetzt es gilt

$$\lim_{n \rightarrow \infty} \|P_n e_k - e_k\|_\infty = 0, \quad 0 \leq k \leq 2, \quad (3.1)$$

für die Monomfunktionen $e_k : x \mapsto x^k$, dann folgt

$$\lim_{n \rightarrow \infty} \|P_n f - f\|_\infty = 0 \quad (3.2)$$

für beliebiges $f \in C([a, b]; \mathbb{R})$.

Beweis. Zunächst einmal notieren wir, dass aus der Voraussetzung (3.1) aufgrund der Linearität der Operatoren P_n schon

$$\lim_{n \rightarrow \infty} \|P_n \psi_\xi\|_\infty = 0 \quad (3.3)$$

für alle $\xi \in [a, b]$ folgt, wobei die Funktionen ψ_ξ durch $\psi_\xi = \psi_\xi(x) = (x - \xi)^2$ definiert sind. Ist also $f \in C([a, b]; \mathbb{R})$, dann gibt es ein $M > 0$ mit $-M < f(x) < M$ für alle $x \in [a, b]$ und es folgt

$$-2M < f(x) - f(\xi) < 2M, \quad x \in [a, b]. \quad (3.4)$$

Aufgrund der Stetigkeit von f bei ξ existiert zu $\varepsilon > 0$ ein $\delta > 0$ mit

$$|x - \xi| < \delta \implies |f(x) - f(\xi)| < \varepsilon, \quad (3.5)$$

so dass mit (3.4) und (3.5) schon

$$-\varepsilon - \frac{2M}{\delta^2} \psi_\xi(x) \leq f(x) - f(\xi) \leq \varepsilon + \frac{2M}{\delta^2} \psi_\xi(x) \quad (3.6)$$

für alle $x \in [a, b]$ folgt; denn ist $|x - \xi| < \delta$ gilt (3.5) und für $|x - \xi| \geq \delta$ ist

$$\frac{\psi_\xi(x)}{\delta^2} \geq 1,$$

womit (3.6) resultiert. Nun greift die Positivität, das heißt die Monotonie der Operatoren P_n in das Geschehen ein. Die Anwendung des monotonen Operators P_n auf Gleichung (3.6) liefert

$$-\varepsilon P_n e_0 - \frac{2M}{\delta^2} P_n \psi_\xi \leq P_n f - f(\xi) P_n e_0 \leq \varepsilon P_n e_0 + \frac{2M}{\delta^2} P_n \psi_\xi. \quad (3.7)$$

Setzen wir in (3.7) $x = \xi$, so folgt aufgrund von (3.3)

$$\lim_{n \rightarrow \infty} |P_n f(\xi) - f(\xi)| \leq \varepsilon.$$

Da $\varepsilon > 0$ unabhängig von ξ gewählt wurde, liefert dies die Behauptung des Satzes. \square

Der Konvergenzsatz von Bohman und Korovkin liefert einen eleganten Beweis des Weierstraßschen Approximationssatzes unter Verwendung von *Bernstein-Polynomen*.

Definition 3.2 Sei $f \in C([0, 1]; \mathbb{R})$ und $n \in \mathbb{N}$. Dann ist durch

$$\mathcal{B}_n f(x) = \sum_{i=0}^n \binom{n}{i} x^i (1-x)^{n-i} f\left(\frac{i}{n}\right), \quad x \in [0, 1],$$

das n -te Bernstein-Polynom zu f gegeben.

Wir sehen sofort, dass durch $\mathcal{B}_n : C([0, 1]; \mathbb{R}) \rightarrow C([0, 1]; \mathbb{R})$ ein linearer und positiver Operator definiert ist. Daher können wir Theorem 3.1 anwenden:

Satz 3.1 (Weierstraßscher Approximationssatz auf $[0, 1]$)

Zu jeder stetigen Funktion $f \in C([0, 1]; \mathbb{R})$ existiert eine Folge $(p_n)_{n \geq 0}$ von Polynomfunktionen $p_n : [0, 1] \rightarrow \mathbb{R}$ mit $\text{grad } p_n = n$, die gleichmäßig auf dem Intervall $[0, 1]$ mit $n \rightarrow \infty$ gegen f konvergieren, das heißt es ist

$$\lim_{n \rightarrow \infty} \|f - p_n\|_\infty = 0.$$

Beweis. Offensichtlich ist durch $\mathcal{B}_n : C([0, 1]; \mathbb{R}) \rightarrow C([0, 1]; \mathbb{R})$ ein sowohl positiver als auch linearer Operator gegeben. Wählen wir $p_n = \mathcal{B}_n f$ und können

$$\lim_{n \rightarrow \infty} \|e_k - \mathcal{B}_n e_k\|_\infty = 0, \quad k = 0, 1, 2, \quad (3.8)$$

für die Monomfunktionen $e_k : x \mapsto x^k$ beweisen, so folgt die Aussage des Satzes direkt durch Anwendung von Theorem 3.1. Für $x \in [0, 1]$ und $n \in \mathbb{N}$ gilt

$$\begin{aligned} \mathcal{B}_n e_0(x) &= \sum_{i=0}^n \binom{n}{i} x^i (1-x)^{n-i} = (x + 1 - x)^n = 1, \\ \mathcal{B}_n e_1(x) &= \sum_{i=0}^n \binom{n}{i} x^i (1-x)^{n-i} \frac{i}{n} = \sum_{i=1}^n \binom{n-1}{i-1} x^i (1-x)^{n-i} \\ &= \sum_{i=0}^{n-1} \binom{n-1}{i} x^{i+1} (1-x)^{n-i-1} \frac{i}{n} = x \mathcal{B}_n e_0(x) = x, \end{aligned}$$

sowie

$$\begin{aligned} \mathcal{B}_n e_2(x) &= \sum_{i=0}^n \binom{n}{i} x^i (1-x)^{n-i} \frac{i^2}{n^2} = \sum_{i=1}^n \binom{n-1}{i-1} x^i (1-x)^{n-i} \frac{i}{n} \\ &= \frac{x}{n} \sum_{i=0}^{n-1} \binom{n-1}{i} x^i (1-x)^{n-i-1} (i+1) \\ &= \frac{x}{n} \left((n-1) \mathcal{B}_{n-1} e_1(x) + \mathcal{B}_{n-1} e_0(x) \right) = x^2 + \frac{1}{n} (x - x^2), \end{aligned}$$

das heißt $\mathcal{B}_n e_2 \rightarrow e_2$ gleichmäßig auf $[0, 1]$ mit $n \rightarrow \infty$. Es folgt unmittelbar die Gültigkeit von Gleichung (3.8). \square

Es sei angemerkt, dass der Weierstraßsche Approximationssatz auch mit Hilfe der Bernstein-Polynome ohne Anwendung des Satzes von Bohman und Korovkin bewiesen werden kann. Die Struktur des obigen Beweises gibt uns aber einen Eindruck davon, wie durch Anwendung von Theorem 3.1 ein Konvergenzbeweis für exakte kleinste Quadrate Interpolierende zu etablieren ist. Hierauf wollen wir im Folgenden eingehen.

3.1.ii Ein reines Konvergenzresultat für die univariate Shepard-Interpolation

Wir hatten in der Definition 1.2 die Shepard-Interpolierende $S_n^\alpha f$ zu einer Funktion $f \in C(K; \mathbb{R})$ kennengelernt. Diese besitzt nach Konstruktion die Interpolationseigenschaft $S_n^\alpha f(x_i) = f(x_i)$ für $x_i \in \{x_1, \dots, x_n\} \subset K$. Im Fall $n \rightarrow \infty$ stellt sich die Frage nach dem Verhalten der Shepard-Interpolierenden $S_n^\alpha f$ zur Funktion f , wobei wir die Interpretation des Grenzwertes $n \rightarrow \infty$ für die Menge $\{x_1, \dots, x_n\}$ relativ zum Intervall $[a, b]$ später näher erläutern werden.

Anstatt eine direkte Konvergenzaussage für beliebige $f \in C(K; \mathbb{R})$ zu beweisen, betrachten wir zunächst die Struktur des Shepard-Interpolationsoperators. Dieser ist positiv im Sinne der Definition 3.1.

Lemma 3.1 Seien $\{x_1, \dots, x_n\} \subset K$ sowie $w_\alpha(x) = |x|^{-\alpha}$ mit einer natürlichen Zahl $\alpha > 1$ fest gewählt. Dann ist der Shepard-Interpolationsoperator $S_n^\alpha : C(K; \mathbb{R}) \rightarrow C(K; \mathbb{R})$ gemäß der Definition 1.2 positiv bezüglich der natürlichen \leq -Ordnungsrelation auf $C(K; \mathbb{R})$.

Beweis. Die natürliche Ordnungsrelation \leq auf $C(K; \mathbb{R})$ ist durch

$$g \leq h \iff \forall x \in K : g(x) \leq h(x)$$

gegeben. Satz 1.4(b) impliziert

$$0 \leq \varphi_i(x) = \frac{w_\alpha(x - x_i)}{\sum_{j=1}^n w_\alpha(x - x_j)}$$

und aus $0 \leq f$ für $f \in C(K; \mathbb{R})$ folgt

$$0 \leq \sum_{i=1}^n f(x_i) \varphi_i(x), \quad x \in K.$$

Damit ist der Operator S_n^α positiv. □

Mit Hilfe von Theorem 3.1 sind wir jetzt in der Lage eine erste reine Konvergenzaussage für die univariate Shepard-Interpolation zu beweisen. Wir betrachten den Fall $K = [a, b]$.

Satz 3.2 Für jedes $n \in \mathbb{N}$ mit $n \geq 2$ sei $a = x_1^{(n)} < \dots < x_{c(n)}^{(n)} = b$ eine Verteilung auf dem Intervall $[a, b]$ mit

$$r_n = \max_{1 \leq i \leq c(n)-1} |x_{i+1}^{(n)} - x_i^{(n)}|.$$

Weiter existiere eine Abbildung $\eta : \mathbb{N} \rightarrow \mathbb{R}$ mit $\lim_{n \rightarrow \infty} \eta(n) = 0$ und

$$c(n) r_n^2 \leq \eta(n).$$

Zu $f \in C([a, b]; \mathbb{R})$ sei $S_n^\alpha f$ die Shepard-Interpolierende zu f gemäß Definition 1.2 bezüglich der obigen Verteilung. Dann konvergiert $S_n^\alpha f$ gleichmäßig gegen f auf dem Intervall $[a, b]$ mit $n \rightarrow \infty$, das heißt

$$\lim_{n \rightarrow \infty} \|S_n^\alpha f - f\|_\infty = 0.$$

Beweis. Nach Lemma 3.1 sowie Theorem 3.1 genügt es zu zeigen, dass

$$\lim_{n \rightarrow \infty} \|S_n^\alpha e_k - e_k\|_\infty = 0, \quad k = 0, 1, 2,$$

für die Monomfunktionen $e_k : x \mapsto x^k$ erfüllt ist.

Die Funktion e_0 ist invariant unter Anwendung des Operators S_n^α nach dessen Konstruktion. Das heißt, es gilt $S_n^\alpha e_0 = e_0$ für alle $n \in \mathbb{N}$ unabhängig von der Wahl der $x_i^{(n)}$ aufgrund von Satz 1.4.

Betrachten wir nun den Fall $k = 1$. Für beliebiges $x \in [a, b]$ ist

$$\begin{aligned}
& |\mathcal{S}_n^\alpha e_1(x) - e_1(x)| \\
& \leq \sum_{j=1}^{c(n)} |x_j^{(n)} - x| \frac{|x - x_j^{(n)}|^{-\alpha}}{\sum_{l=1}^{c(n)} |x - x_l^{(n)}|^{-\alpha}} \\
& = \sum_{\substack{j \\ |x - x_j^{(n)}| < r_n^2}} |x_j^{(n)} - x| \frac{|x - x_j^{(n)}|^{-\alpha}}{\sum_{l=1}^{c(n)} |x - x_l^{(n)}|^{-\alpha}} \\
& \quad + \sum_{\substack{j \\ |x - x_j^{(n)}| \geq r_n^2}} |x_j^{(n)} - x| \frac{|x - x_j^{(n)}|^{-\alpha}}{\sum_{l=1}^{c(n)} |x - x_l^{(n)}|^{-\alpha}} \\
& \leq r_n^2 + \sum_{\substack{j \\ |x - x_j^{(n)}| \geq r_n^2}} |x_j^{(n)} - x| \frac{|x - x_j^{(n)}|^{-\alpha}}{\sum_{l=1}^{c(n)} |x - x_l^{(n)}|^{-\alpha}} \\
& = r_n^2 + c(n) \mathcal{O}(r_n^2) = \mathcal{O}(r_n^2 + \eta(n)),
\end{aligned}$$

aufgrund von

$$|x - x_j^{(n)}|^{-\alpha+1} = \mathcal{O}(r_n^{2(-\alpha+1)})$$

und

$$\sum_{l=1}^{c(n)} |x - x_l^{(n)}|^{-\alpha} \geq \sum_{\substack{l \\ |x - x_l^{(n)}| < r_n^2}} |x - x_l^{(n)}|^{-\alpha} \geq r_n^{-2\alpha}.$$

Schließlich sei $k = 2$ und wir setzen $M = \max_{x \in [a,b]} |x|$. Dann erhalten wir

$$\begin{aligned}
|\mathcal{S}_n^\alpha e_2(x) - e_2(x)| & \leq \sum_{j=1}^{c(n)} |(x_j^{(n)})^2 - x^2| \frac{|x - x_j^{(n)}|^{-\alpha}}{\sum_{l=1}^{c(n)} |x - x_l^{(n)}|^{-\alpha}} \\
& \leq 2M \sum_{j=1}^{c(n)} |x_j^{(n)} - x| \frac{|x - x_j^{(n)}|^{-\alpha}}{\sum_{l=1}^{c(n)} |x - x_l^{(n)}|^{-\alpha}} \\
& = \mathcal{O}(r_n^2 + \eta(n)),
\end{aligned}$$

durch Betrachtung des Falls $k = 1$. Insgesamt liefern die obigen Resultate für $k = 0, 1, 2$ schon

$$\|\mathcal{S}_n^\alpha e_k - e_k\|_\infty = \sup_{x \in [a,b]} |\mathcal{S}_n^\alpha e_k(x) - e_k(x)| = \mathcal{O}(r_n^2 + \eta(n)),$$

woraus mit $n \rightarrow \infty$ unmittelbar die zu beweisende Konvergenz

$$\lim_{n \rightarrow \infty} \|\mathcal{S}_n^\alpha e_k - e_k\|_\infty = 0, \quad k = 0, 1, 2,$$

folgt. □

3.1.iii Ein Konvergenzresultat unter Anwendung von Glattheitsmoduln

Auf der einen Seite gibt Satz 3.2 eine reine Konvergenzaussage für die univariate Shepard-Interpolation, auf der anderen Seite lässt dieser allerdings die Frage nach der Güte der einzelnen Approximation $S_n^\alpha f$ offen. Diese Frage beantworten wir innerhalb dieses Abschnittes mit Hilfe von Glattheitsmoduln, die im Folgenden für stetige reellwertige Funktionen auf einem Intervall $[a, b] \subset \mathbb{R}$ eingeführt werden. Wir orientieren uns hierbei an Altomare und Campiti [3].

Definition 3.3 Sei $f \in C([a, b]; \mathbb{R})$ eine stetige reellwertige Funktion, $x \in [a, b]$, $h \in \mathbb{R}$ mit $h \neq 0$ und $k \in \mathbb{N}$. Dann heißt

$$\Delta_h^k f(x) = \sum_{m=0}^k (-1)^{m+k} \binom{k}{m} f(x + mh)$$

die k -te Differenz von f bei x mit Schrittweite h , vorausgesetzt dass $x + kh \in [a, b]$ gilt.

Ist weiter $\delta \in \mathbb{R}$ mit $\delta > 0$, so nennt man

$$\omega(f, \delta) = \sup_{\substack{x, y \in [a, b] \\ |x - y| \leq \delta}} |f(x) - f(y)| = \sup_{\substack{|h| \leq \delta \\ x, x+h \in [a, b]}} |\Delta_h f(x)|$$

den *Stetigkeitsmodul* von f mit Argument δ , der auch *Glattheitsmodul erster Ordnung* zu f genannt wird. Allgemeiner ist für $\delta > 0$ der *Glattheitsmodul k -ter Ordnung* zu f durch

$$\omega_k(f, \delta) = \sup_{\substack{|h| \leq \delta \\ x, x+kh \in [a, b]}} |\Delta_h^k f(x)|$$

definiert.

Zur Vereinfachung schreiben wir $\Delta_h f(x) = \Delta_h^1 f(x)$ und die $(k+1)$ -te Differenz berechnet sich rekursiv durch

$$\Delta_h^{k+1} f(x) = \Delta_h^k f(x+h) - \Delta_h^k f(x).$$

Das nachstehende Lemma liefert einige elementare Eigenschaften der Stetigkeits- und Glattheitsmoduln. Wir geben diese Ergebnisse ohne Beweis an und verweisen auf [3].

Lemma 3.2 Sei $f \in C([a, b]; \mathbb{R})$ und $k \in \mathbb{N}$. Dann gelten die folgenden Aussagen:

- (1) Für $0 < \delta_1 \leq \delta_2$ ist $\omega_k(f, \delta_1) \leq \omega_k(f, \delta_2)$.
- (2) Es gilt $\lim_{\delta \rightarrow 0^+} \omega_k(f, \delta) = 0$, falls f gleichmäßig stetig ist.
- (3) Für $\delta > 0$ und $\lambda > 0$ ist $\omega_k(f, \lambda\delta) \leq (1 + [\lambda])^k \omega_k(f, \delta)$, wobei $[\cdot]$ die Gauß-Klammer bezeichnet, das heißt

$$[s] = \max\{n \in \mathbb{N} \mid n \leq s\}, \quad s \in \mathbb{R}_{>0}.$$

Eine Abschätzung von $|S_n^\alpha f(x) - f(x)|$ unter Zuhilfenahme eines Stetigkeitsmoduls für den allgemeinsten Fall, bei dem für die Funktion f lediglich Stetigkeit vorausgesetzt wird, werden wir in Abschnitt 3.7 kennenlernen. An dieser Stelle beschäftigen wir uns mit dem Spezialfall einer differenzierbaren Funktion $f \in C^1([a, b]; \mathbb{R})$. Das weitere Vorgehen orientiert sich an der Monographie von DeVore [6].

Für $x, t \in [a, b]$ gilt die Identität

$$f(t) = f(x) + f'(x)(t - x) + \left(f(t) - f(x) - f'(x)(t - x) \right), \quad (3.9)$$

wobei wir im Fall $x, t \in \{a, b\}$ lediglich einseitige Grenzwerte betrachten wollen. Nehmen wir x als fest gewählt und t als variabel an. Wird der Shepard-Interpolationsoperator \mathcal{S}_n^α auf die Gleichung (3.9) angewendet, liefert dies

$$\mathcal{S}_n^\alpha f(x) - f(x) = f'(x) \mathcal{S}_n^\alpha(t - x)|_{t=x} + \mathcal{S}_n^\alpha \left(f(t) - f(x) - f'(x)(t - x) \right) \Big|_{t=x}, \quad (3.10)$$

wobei $\mathcal{S}_n^\alpha(t - x)|_{t=x}$ in dem Sinne zu interpretieren ist, dass der Operator \mathcal{S}_n^α auf die Funktion $t \mapsto (t - x)$ wirkt und nach dessen Anwendung $t = x$ gesetzt wird. Eine analoge Behandlung gilt für den Term

$$\mathcal{S}_n^\alpha \left(f(t) - f(x) - f'(x)(t - x) \right) \Big|_{t=x}.$$

Eine Abschätzung von (3.10) erhalten wir durch

$$|\mathcal{S}_n^\alpha f(x) - f(x)| \leq |f'(x)| \mathcal{S}_n^\alpha(|t - x|)|_{t=x} + \mathcal{S}_n^\alpha \left(|f(t) - f(x) - f'(x)(t - x)| \right) \Big|_{t=x}, \quad (3.11)$$

denn für den linearen positiven Operator \mathcal{S}_n^α gilt die Ungleichung

$$|\mathcal{S}_n^\alpha f| \leq \mathcal{S}_n^\alpha |f|. \quad (3.12)$$

Ist $f \geq 0$ folgt (3.12) aus der Positivität von \mathcal{S}_n^α und für $f < 0$ über die Betrachtung von $-f$ als positive Funktion mit Hilfe der Linearität und Positivität des Operators \mathcal{S}_n^α .

Nun ist aber für ein $t_x \in (\min\{t, x\}, \max\{t, x\})$ und jedes $\delta \in \mathbb{R}$ mit

$$\delta > \max \left\{ |t - x|, \mathcal{S}_n^\alpha((t - x)^2) \Big|_{t=x} \right\}$$

schon

$$\begin{aligned} |f(t) - f(x) - f'(x)(t - x)| &= |f'(t_x) - f'(x)| |t - x| \\ &\leq \omega(f', \delta) |t - x| \\ &\leq \omega(f', \delta) (1 + \delta^{-1} |t - x|) |t - x|, \end{aligned}$$

so dass wir mit Lemma 3.2 die Ungleichung

$$\mathcal{S}_n^\alpha(|f(t) - f(x) - f'(x)(t - x)|) \leq \omega(f', \delta) \left(\mathcal{S}_n^\alpha(|t - x|) + \delta^{-1} \mathcal{S}_n^\alpha((t - x)^2) \right)$$

erhalten. Eine Anwendung dieser Ungleichung bei (3.11) liefert wiederum

$$\begin{aligned} |\mathcal{S}_n^\alpha f(x) - f(x)| &\leq |f'(x)| \mathcal{S}_n^\alpha(|t - x|)|_{t=x} \\ &\quad + \omega(f', \delta) \left(\mathcal{S}_n^\alpha(|t - x|)|_{t=x} + \delta^{-1} \mathcal{S}_n^\alpha((t - x)^2)|_{t=x} \right). \end{aligned} \quad (3.13)$$

Die rechte Seite von (3.13) kann weiter durch eine Ungleichung vom Cauchy-Schwarz-Typ für positive lineare Operatoren der Form

$$\mathcal{S}_n^\alpha(|t - x|)|_{t=x} \leq \sqrt{\mathcal{S}_n^\alpha((t - x)^2)|_{t=x}}$$

abgeschätzt werden, so dass wir die Güte der Approximation von $\mathcal{S}_n^\alpha f$ an f im Wesentlichen durch die Güte der Approximation von $\mathcal{S}_n^\alpha e_k$ an e_k für $0 \leq k \leq 2$ messen können.

Wir kommen im weiteren Verlauf auf die Cauchy-Schwarz-Ungleichung zurück und verweisen zunächst auf [15].

Die Abschnitte 3.1.ii sowie 3.1.iii stellen eine Art roten Faden für weitere Teile dieser Arbeit dar. Wir werden im Weiteren das Konzept von *fast positiven linearen Operatoren* auf dem Raum $C(K; \mathbb{R})$ einführen und für diese Folgen eine multivariate Erweiterung von Satz 3.1 beweisen. Darauf aufbauend liefern wir für gewisse monomiale Interpolationsbasen nach Definition 1.5 einen Konvergenzbeweis für exakte kleinste Quadrate Interpolierende in Analogie zu Satz 3.2. Dieser Beweis verläuft wiederum ähnlich den Erläuterungen des Abschnitts 3.1.iii.

3.2 Über die Namensgebung des Konvergenzsatzes von Bohman und Korovkin

Wir wollen kurz auf die Entstehungsgeschichte von Theorem 3.1, das heißt dem klassischen Konvergenzsatz von Bohman und Korovkin, eingehen. Anders als es die Namensgebung vermuten lässt, war es der rumänische Mathematiker Tiberiu Popoviciu, der ein erstes Resultat ähnlich zu dem uns bekannten Theorem 3.1 formulierte und bewies [21]. Da es in rumänischer Sprache veröffentlicht wurde, war es über einen langen Zeitraum dem größeren Fachpublikum verborgen. Mehrere Jahrzehnte nach Veröffentlichung Popovicius Resultat, wurde durch Kascó [22] eine in die englische Sprache übersetzte Version von [21] der Öffentlichkeit zugänglich.

Popoviciu ging seinerzeit der Frage nach, wie man eine trianguläre Matrix von Zahlen aus dem Intervall $[a, b]$ der Form

$$x_{n,0}, x_{n,1}, \dots, x_{n,n}, \quad n = 1, 2, 3, \dots,$$

und eine trianguläre Matrix von Polynomen $P_{n,m} = P_{n,m}(x)$ der Form

$$P_{n,0}, P_{n,1}, \dots, P_{n,n}, \quad n = 1, 2, 3, \dots,$$

zu wählen hat, so dass zu beliebigem $f \in C([a, b]; \mathbb{R})$ mit

$$Q_n f(x) = \sum_{i=0}^n P_{n,i}(x) f(x_{n,i}), \quad n = 1, 2, 3, \dots \quad (3.14)$$

schon

$$\lim_{n \rightarrow \infty} Q_n f(x) = f(x), \quad x \in [a, b], \quad (3.15)$$

erfüllt ist. Er konnte zeigen, sofern

$$P_{n,i}(x) \geq 0, \quad i = 0, 1, \dots, n, \quad n = 1, 2, 3, \dots, \quad x \in [a, b], \quad (3.16a)$$

$$\sum_{i=0}^n P_{n,i}(x) = 1, \quad n = 1, 2, 3, \dots, \quad (3.16b)$$

gelten, dass die Differenz $Q_n f(x) - f(x)$ durch

$$|Q_n f(x) - f(x)| \leq 2\omega(f, A_n)$$

mit

$$A_n = \sup_{x \in [a,b]} \sum_{i=0}^n |x_{n,i} - x| P_{n,i}(x)$$

abgeschätzt werden kann. Dieses Resultat ermöglichte es ihm für $[a, b] = [0, 1]$ durch das Setzen von $x_{n,i} = \frac{i}{n}$ sowie $P_{n,i}(x) = \binom{n}{i} x^i (1-x)^{n-i}$, den Weierstraßschen Approximationssatz zu beweisen. Hierbei erkennen wir sofort, dass die Bedingungen (3.16a) und (3.16b) durch diese spezielle Wahl von $x_{n,i}$ und $P_{n,i}$ erfüllt sind.

Harald Bohman [5] veröffentlichte nahezu zeitgleich zu Popoviciu Konvergenzuntersuchungen für Operatoren, die ebenfalls von der Form (3.14) sind. Im Gegensatz zu Popoviciu verzichtete er allerdings bei der Wahl der $P_{n,i}$ auf die Einschränkung auf Polynomfunktionen und betrachtete beliebige nichtnegative Funktionen. Bohman zeigte hierbei hinreichende und notwendige Bedingungen an die Punkte $x_{n,i}$ und Funktionen $P_{n,i}$ in einem Kontext der Konvergenzfrage (3.15) auf.

Wenige Jahre später brachte Pavel Petrovich Korovkin ein Konvergenzresultat über lineare und positive Operatoren auf Räumen stetiger Funktionen [14] zunächst im Kontext von Integraloperatoren (vgl. [2]). In der 1960 erschienenen Monographie [15] von Korovkin findet sich zum ersten Mal eine Version des Konvergenzsatzes in der uns bekannten Form entsprechend Theorem 3.1.

Aus heutiger Sicht scheint speziell die Tatsache, dass Popoviciu sein Resultat in rumänischer Sprache veröffentlicht hat, ein Hauptgrund dafür zu sein, dass Konvergenzresultate im Stil von Theorem 3.1 in der Approximationstheorie unter den Namen Bohman und Korovkin Einzug gehalten haben.

3.3 Konvergenz von Folgen fast positiver linearer Operatoren auf $C(K; \mathbb{R})$

Dieser Abschnitt führt den Begriff einer *Folge von fast positiven linearen Operatoren* auf Räumen $C(K; \mathbb{R})$ als spezielle Folge von regulären Operatoren ein. Diese Begriffsbildung ermöglicht die Verallgemeinerung einer Konvergenzaussage vom Korovkin-Typ für Folgen von positiven Operatoren. Im darauf folgenden Teil dieser Arbeit wenden wir dieses Resultat zum Beweis der Konvergenz von exakte kleinste Quadrate Interpolierenden an.

Definition 3.4 Seien $\Phi, \Phi_n : C(K; \mathbb{R}) \rightarrow C(K; \mathbb{R})$ für $n \geq n_0 \in \mathbb{N}$ lineare Operatoren.

(a) Der Operator Φ wird *positiv* genannt, falls

$$\Phi \varphi \geq 0$$

für alle nichtnegativen $\varphi \in C(K; \mathbb{R})$ gilt.

(b) Der Operator Φ_n heißt *regulär*, falls positive lineare Operatoren

$$\Phi_n^+, \Phi_n^- : C(K; \mathbb{R}) \rightarrow C(K; \mathbb{R})$$

mit der Eigenschaft

$$\Phi_n = \Phi_n^+ - \Phi_n^- \quad (3.17)$$

existieren.

(c) Eine Folge $(\Phi_n)_{n \geq n_0}$ von linearen Operatoren

$$\Phi_n : C(K; \mathbb{R}) \rightarrow C(K; \mathbb{R})$$

wird *fast positiv* genannt, falls jedes Φ_n regulär mit im Grenzfalle $n \rightarrow \infty$ verschwindendem Negativteil Φ_n^- ist, das heißt es existieren zwei Folgen von positiven linearen Operatoren $(\Phi_n^+)_{n \geq n_0}$ und $(\Phi_n^-)_{n \geq n_0}$ in $C(K; \mathbb{R})$ mit der Eigenschaft (3.17) für alle $n \geq n_0$ und es gilt

$$\lim_{n \rightarrow \infty} \Phi_n^- \varphi = 0$$

für alle $\varphi \in C(K; \mathbb{R})$.

Die Darstellung (3.17) ist dabei nicht notwendigerweise eindeutig.

Wir hatten in Satz 3.2 festgestellt, dass die Konvergenz des positiven Operators P_n für die Monomfunktionen e_0, e_1 und e_2 schon die Konvergenz für beliebiges $f \in C([a, b]; \mathbb{R})$ impliziert. Eine Menge von Funktionen mit dieser speziellen Eigenschaft wird *Korovkin-Menge* für eine gewisse Operatorfolge genannt.

Definition 3.5 Sei $(\Phi_n)_{n \geq n_0}$ eine Folge von Operatoren, die den Raum $C(K; \mathbb{R})$ in sich selbst abbilden und $\mathcal{K} \subset C(K; \mathbb{R})$ eine Menge von Funktionen mit der Eigenschaft

$$\lim_{n \rightarrow \infty} \|h - \Phi_n h\|_\infty = 0 \quad \forall h \in \mathcal{K} \quad \implies \quad \lim_{n \rightarrow \infty} \|f - \Phi_n f\|_\infty = 0 \quad \forall f \in C(K; \mathbb{R}).$$

Dann wird \mathcal{K} eine $id_{C(K; \mathbb{R})}$ -Korovkin-Menge für $(\Phi_n)_{n \geq n_0}$ genannt.

Die Definition 3.5 kann für beliebige Operatoren Φ in der Form „ $\lim_{n \rightarrow \infty} \|\Phi h - \Phi_n h\|_\infty = 0$ für alle $h \in \mathcal{K}$ impliziert $\lim_{n \rightarrow \infty} \|\Phi f - \Phi_n f\|_\infty = 0$ für beliebiges $f \in C(K; \mathbb{R})$ “ verallgemeinert werden. Wir beschränken uns aber auf den Fall $\Phi = id_{C(K; \mathbb{R})}$, der die natürliche Wahl für spätere Anwendungen dieser Arbeit darstellt. Für Betrachtungen mit allgemeinem Operator Φ verweisen wir auf die Monographie [3] von Altomare und Campiti.

Definition 3.6 Die Funktionen $f_1, \dots, f_m \in C(K; \mathbb{R})$ mit $f_1 = e_0$ werden *positiv summierbar* genannt, falls die folgenden Eigenschaften gelten:

(1) Es existieren Funktionen $a_1, \dots, a_m \in C(K; \mathbb{R})$, so dass der Ausdruck

$$P_y(x) = \sum_{i=1}^m a_i(y) f_i(x)$$

nichtnegativ für alle $x, y \in K$ ist.

(2) Es gibt Skalare $\alpha_1, \dots, \alpha_m \in \mathbb{R}$ derart, dass

$$\sum_{i=1}^m \alpha_i f_i(x)$$

positiv für alle $x \in K$ ist.

(3) Für $x, y \in K$ ist $P_y(x) = 0$ genau dann wenn $x = y$ gilt.

Zur Motivation der gerade gegebenen Definition von *positiv summierbaren Funktionen* schauen wir uns im Zusammenhang mit den Korovkin-Mengen ein Beispiel an: Die Monomfunktionen e_0, e_1 und e_2 aus Theorem 3.1 sind positiv summierbar. Denn nehmen wir $f_1(x) = e_0(x) = 1$, $f_2(x) = e_1(x) = x$ und $f_3(x) = e_2(x) = x^2$ an, liefert das Setzen von $a_1(y) = y^2$, $a_2(y) = -2y$ sowie $a_3(y) = 1$ schon, dass

$$P_y(x) = \sum_{i=1}^3 a_i(y) f_i(x) = (x - y)^2$$

nichtnegativ ist. Ferner gilt $P_y(x) = 0 \iff x = y$ und die letzte zu erfüllende Bedingung erreichen wir durch das Setzen von $\alpha_1 = \alpha_2 = 2$ sowie $\alpha_3 = 1$.

An dieser Stelle kann vermutet werden, dass Mengen von positiv summierbaren Funktionen Prototypen der Korovkin-Mengen für gewisse Folgen von Operatoren nach Definition 3.5 darstellen. Diese Vermutung stellt sich in der Tat als richtig heraus, wie uns Theorem 3.2 als Hauptresultat dieses Kapitels lehrt.

Theorem 3.2 Sei $(\Phi_n)_{n \geq n_0}$ eine Folge von fast positiven linearen Operatoren, die den Raum $C(K; \mathbb{R})$ in sich selbst abbilden. Für jedes $n \geq n_0$ existiere eine Darstellung

$$\Phi_n = \Phi_n^+ - \Phi_n^-,$$

wobei durch $\Phi_n^+, \Phi_n^- : C(K; \mathbb{R}) \rightarrow C(K; \mathbb{R})$ positive lineare Operatoren gegeben sind. Darüber hinaus sei $\mathcal{K} = \{f_1, \dots, f_m\}$ eine Menge von positiv summierbaren Funktionen aus $C(K; \mathbb{R})$ mit der Eigenschaft

$$\lim_{n \rightarrow \infty} \|h - \Phi_n h\|_\infty = 0 \quad \forall h \in \mathcal{K}.$$

Dann ist \mathcal{K} eine $id_{C(K; \mathbb{R})}$ -Korovkin-Menge für $(\Phi_n)_{n \geq n_0}$, das heißt es gilt

$$\lim_{n \rightarrow \infty} \|f - \Phi_n f\|_\infty = 0$$

für beliebiges $f \in C(K; \mathbb{R})$.

Das Theorem 3.2 ist die direkte Verallgemeinerung eines Resultates für Folgen von positiven Operatoren, welches bei Lorentz [19] in einer klassischen Darstellung zu finden ist. Für den Beweis von Theorem 3.2 geben wir im weiteren Verlauf einige Hilfsresultate an. Lemma 3.3 befindet sich in der hier dargestellten Form in Lorentz' Monographie für Folgen von positiven Operatoren. Wir geben dieses Ergebnis dennoch samt Beweis aus Gründen der Vollständigkeit an.

Lemma 3.3 Sei $(\Phi_n)_{n \geq n_0}$ eine Folge von fast positiven linearen Operatoren und $\mathcal{K} = \{f_1, \dots, f_m\}$ eine Menge von positiv summierbaren Funktionen wie in Theorem 3.2. Ferner sei $P_y(x)$ eine Darstellung für die Funktionen in \mathcal{K} nach Definition 3.6. Dann gelten die folgenden Aussagen:

- (1) Für jedes $y \in K$ ist $\lim_{n \rightarrow \infty} \Phi_n P_y(x) \Big|_{x=y} = 0$.
- (2) Sind die Funktionen $a_1(y) = a_1, \dots, a_m(y) = a_m$ konstant, das heißt

$$P(x) \equiv P_y(x) = \sum_{i=1}^m a_i f_i(x),$$

so gilt $\lim_{n \rightarrow \infty} \|P - \Phi_n P\|_\infty = 0$.

Beweis. Zunächst einmal folgt (1) aus

$$\Phi_n P_y(x) = \sum_{i=1}^m a_i(y) \Phi_n f_i(x) \longrightarrow \sum_{i=1}^m a_i(y) f_i(x), \quad n \rightarrow \infty,$$

unter Beachtung von $P_y(x) = 0 \iff x = y$ nach Definition 3.6. Zum anderen ist (2) eine direkte Konsequenz von

$$\|P - \Phi_n P\|_\infty = \left\| \sum_{i=1}^m a_i (f_i - \Phi_n f_i) \right\|_\infty \leq \max_{i=1, \dots, m} |a_i| \cdot \sum_{i=1}^m \|f_i - \Phi_n f_i\|_\infty \longrightarrow 0$$

mit $n \rightarrow \infty$. □

Den Beweis des folgenden Lemmas haben wir ausgehend von der Originaldarstellung in [19] derart angepasst, dass das Originalresultat für den positiven Operator

$$\Phi_n^+ = \Phi_n + \Phi_n^-$$

angewendet wird. Unser Ergebnis folgt im Grenzfall für $n \rightarrow \infty$ aufgrund des verschwindenden Negativteils Φ_n^- .

Lemma 3.4 Sei $\{f_y \mid y \in K\}$ eine K -indizierte Familie von stetigen Funktionen $f_y \in C(K; \mathbb{R})$ derart, dass die Abbildung

$$K \times K \ni (x, y) \mapsto f_y(x) \in \mathbb{R} \tag{3.18}$$

stetig von sowohl x als auch y abhängt und $f_y(y) = 0$ für alle $y \in K$ gilt. Durch $(\Phi_n)_{n \geq n_0}$ sei eine Folge von fast positiven linearen Operatoren wie in Theorem 3.2 gegeben. Dann existiert der Grenzwert

$$\lim_{n \rightarrow \infty} \Phi_n f_y(x) \Big|_{x=y} = 0$$

gleichmäßig in y .

Beweis. Sei $\varepsilon > 0$ und $\mathcal{D} := \{(y, y) \mid y \in K\} \subset K \times K$. Aufgrund der vorausgesetzten Stetigkeitsbedingung (3.18) gibt es für jedes $(y, y) \in \mathcal{D}$ eine offene Umgebung U_y mit

$$(x, y) \in U_y \iff |f_y(x)| < \varepsilon.$$

Es bezeichne V die Vereinigung aller solcher offenen Umgebungen und sei

$$E = (K \times K) \setminus V$$

deren Komplement in $K \times K$. Offensichtlich ist die Menge E kompakt. Sei nun $P_y(x)$ wie in Lemma 3.3 und seien $m, M \in \mathbb{R}$ durch

$$m := \min_{(x,y) \in E} P_y(x), \quad M := \max_{(x,y) \in E} |f_y(x)|$$

definiert. Dann ist für beliebiges $(x, y) \in K \times K$

$$|f_y(x)| \leq \varepsilon \cdot f_1(x) + \frac{M}{m} P_y(x)$$

oder alternativ

$$-\varepsilon \cdot f_1(x) - \frac{M}{m} P_y(x) \leq f_y(x) \leq \varepsilon \cdot f_1(x) + \frac{M}{m} P_y(x). \quad (3.19)$$

Bis zu dieser Stelle haben wir den Beweis des Lemmas, wie bei Lorentz [19] dargestellt, übernommen. Erinnern wir uns nun an die Äquivalenz von Positivität und Monotonie von linearen Operatoren. Eine Anwendung des positiven linearen Operators $\Phi_n^+ = \Phi_n + \Phi_n^-$ auf Gleichung (3.19) liefert dann

$$\begin{aligned} -\varepsilon \cdot (\Phi_n f_1(x) + \Phi_n^- f_1(x)) - \frac{M}{m} (\Phi_n P_y(x) + \Phi_n^- P_y(x)) &\leq \Phi_n f_y(x) + \Phi_n^- f_y(x) \\ &\leq \varepsilon \cdot (\Phi_n f_1(x) + \Phi_n^- f_1(x)) + \frac{M}{m} (\Phi_n P_y(x) + \Phi_n^- P_y(x)), \end{aligned}$$

oder in äquivalenter Schreibweise

$$|\Phi_n f_y(x) + \Phi_n^- f_y(x)| \leq \varepsilon \cdot (\Phi_n f_1(x) + \Phi_n^- f_1(x)) + \frac{M}{m} (\Phi_n P_y(x) + \Phi_n^- P_y(x)).$$

Das heißt für festes $x \in K$ gilt

$$\left| \lim_{n \rightarrow \infty} \Phi_n f_y(x) \right| \leq \varepsilon \cdot \lim_{n \rightarrow \infty} \Phi_n f_1(x) + \frac{M}{m} \lim_{n \rightarrow \infty} \Phi_n P_y(x), \quad (3.20)$$

da die Φ_n^- -Anteile der Operatoren Φ_n mit $n \rightarrow \infty$ entsprechend ihrer Definition verschwinden. Setzen wir $x = y$ in Gleichung (3.20), so liefert dies

$$\left| \lim_{n \rightarrow \infty} \Phi_n f_y(x) \right|_{x=y} \leq \varepsilon \cdot \lim_{n \rightarrow \infty} \Phi_n f_1(y) = \varepsilon$$

aufgrund von Lemma 3.3. Das heißt der Grenzwert $\lim_{n \rightarrow \infty} \Phi_n f_y(x) \Big|_{x=y} = 0$ existiert gleichmäßig in $y \in K$. \square

Jetzt sind alle Hilfsmittel bereitgestellt, die für den Beweis von Theorem 3.2 benötigt werden.

Beweis von Theorem 3.2. Sei $f \in C(K; \mathbb{R})$ und sei $\{f_y \mid y \in K\}$ die K -indizierte Familie von Funktionen definiert durch

$$f_y(x) := f(x) - \frac{f(y)}{P(y)} P(x),$$

wobei die Koeffizienten α_i im Ausdruck $P = P(y) = \sum_{i=1}^m \alpha_i f_i(y)$ nach Definition 3.6 derart zu wählen sind, dass $P(y) > 0$ für jedes $y \in K$ gilt. Dann hängt die Abbildung

$$K \times K \ni (x, y) \mapsto f_y(x) \in \mathbb{R}$$

stetig von x sowie y ab. Darüber hinaus gilt gleichmäßig für $y \in K$ die Gleichung

$$\begin{aligned} 0 = \lim_{n \rightarrow \infty} \Phi_n f_y(x) \Big|_{x=y} &= \lim_{n \rightarrow \infty} \Phi_n f(x) \Big|_{x=y} - \frac{f(y)}{P(y)} \lim_{n \rightarrow \infty} \Phi_n P(x) \Big|_{y=x} \\ &= \lim_{n \rightarrow \infty} \Phi_n f(x) \Big|_{x=y} - f(y) \end{aligned}$$

nach Lemma 3.3 und Lemma 3.4. □

3.4 Das Konvergenzproblem für die exakte kleinste Quadrate Interpolation

Wir erläutern im Folgenden das Konvergenzproblem für die exakte kleinste Quadrate Interpolation. Hierfür entwickeln wir den Begriff einer *monotonen Folge von Interpolationsknotenverteilungen*. Diese Begriffsbildung und die Resultate in Abschnitt 3.3 bilden die Basis der Konvergenzbeweise für exakte kleinste Quadrate Interpolierende in den Abschnitten 3.5 und 3.7.

Bisher haben wir lediglich bezüglich einer fest gewählten Menge $\{x_1, \dots, x_n\} \subset K$ für eine Funktion $f \in C(K; \mathbb{R})$ die exakte kleinste Quadrate Interpolierende $\mathcal{G}f : K \rightarrow \mathbb{R}$ mit

$$\mathcal{G}f(x_i) = f(x_i)$$

für $1 \leq i \leq n$ bestimmt. Konvergenzaussagen betreffen den Fall $n \rightarrow \infty$. Wir benötigen daher eine Definition des Ausdrucks $n \rightarrow \infty$ für die Mengen $\{x_1, \dots, x_n\}$ im Zusammenhang mit der exakte kleinste Quadrate Interpolation. Die Mengen $\{x_1, \dots, x_n\}$ wollen wir ab sofort als *Interpolationsknotenverteilung* bezeichnen.

Definition 3.7 Sei $n_0 \in \mathbb{N}$ und $\mathcal{B}_{n_0} = \{x_1, \dots, x_{c(n_0)}\} \subset K \subset \mathbb{R}^d$ eine Interpolationsknotenverteilung. Dann wird $(\mathcal{B}_n)_{n \geq n_0}$ eine *monotone Folge von Interpolationsknotenverteilungen* in K genannt, falls eine Abbildung $\eta : \mathbb{N} \rightarrow \mathbb{R}$ mit $\lim_{n \rightarrow \infty} \eta(n) = 0$ existiert, so dass die Eigenschaften

- (1) $\mathcal{B}_n \subset \mathcal{B}_{n+1} \subset K$ für alle $n \geq n_0$,
- (2) $c(n) r(\mathcal{B}_n)^{d+1} \leq \eta(n)$ für $n \geq n_0$ mit

$$r(\mathcal{B}_n) = \inf \{ \delta > 0 \mid \forall x \in K : \text{card}(\mathcal{B}_\delta(x) \cap \mathcal{B}_n) \geq 1 \}$$

erfüllt sind. Durch $c(n) = \text{card} \mathcal{B}_n$ wird die Kardinalität der Menge \mathcal{B}_n angegeben. Ferner bezeichnet $i(n)$ den Index von $x_i = x_{i(n)}$ in \mathcal{B}_n und $\mathcal{B}_\delta(x) = \{y \in \mathbb{R}^d \mid \|x - y\|_2 < \delta\}$.

Beispiel. Sei $K = \{(x_1, \dots, x_d)^T \in \mathbb{R}^d \mid 0 \leq x_i \leq 1 \text{ für } 1 \leq i \leq d\}$ und sei \mathcal{B}_n durch

$$\mathcal{B}_n = K \cap \left\{ \frac{z}{n} \in \mathbb{R}^d \mid z \in \mathbb{Z}^d \right\}$$

für $n \geq 1$ definiert. Dann ist $c(n) = \text{card} \mathcal{B}_n = (n+1)^d$ und

$$r(\mathcal{B}_n) = c_d \frac{1}{n}, \quad (3.21)$$

für $c_d := \sqrt{d \cdot 2^{-d}}$, das heißt

$$r(\mathcal{B}_n)^{d+1} c(n) = (c_d)^{d+1} \frac{1}{n^{d+1}} (n+1)^d = (c_d)^{d+1} \frac{1}{n} \left(1 + \frac{1}{n}\right)^d \leq \underbrace{(c_d)^{d+1} \frac{2^d}{n}}_{=: \eta(n)}.$$

Aufgrund von $\lim_{n \rightarrow \infty} \eta(n) = 0$ ist aber $(\mathcal{B}_n)_{n \geq 2}$ eine monotone Folge von Interpolationsknotenverteilungen in K . Die Gültigkeit der Gleichung (3.21) folgt direkt aus einer geometrischen Überlegung, denn die Euklidische Norm ν des Schwerpunktes des d -dimensionalen Würfels

$$K_n = \{(x_1, \dots, x_d)^T \in \mathbb{R}^d \mid 0 \leq x_i \leq 1/n \text{ für } 1 \leq i \leq d\}$$

hat den Wert $\nu = \frac{\sqrt{d \cdot 2^{-d}}}{n}$.

Bringen wir zunächst noch einmal die Kardinalbasis-Darstellung der exakte kleinste Quadrate Interpolierenden nach Abschnitt 1.6 in Erinnerung: Die Interpolierende ist in dieser durch

$$\mathcal{G}f(x) = \sum_{i=1}^n f(x_{i(n)}) \varphi_i(x)$$

für gewisse Kernfunktionen φ_i nach (1.62) gegeben. Die obiger Konstruktion unterliegende Interpolationsknotenverteilung sei $\mathcal{B}_n = \{x_1, \dots, x_{c(n)}\}$. In einem Fall wie diesem benutzen wir ab sofort die Notation

$$\mathcal{G}_n f(x) = \sum_{i=1}^{c(n)} f(x_{i(n)}) \varphi_i(x).$$

Hierbei verweist n auf die Verteilung \mathcal{B}_n . Im Kontext von Grenzwertbetrachtungen der Art $n \rightarrow \infty$ bei monotonen Folgen von Interpolationsknotenverteilungen erweist sich die Notation \mathcal{G}_n als durchaus nützlich, wie wir im weiteren Verlauf feststellen werden. Weiter bezeichne $\varphi_{i(n)}$ die entsprechende i -te Kernfunktion des zugehörigen Interpolationsoperators \mathcal{G}_n , der bezüglich der Verteilung $\mathcal{B}_n = \{x_1, \dots, x_{c(n)}\}$ konstruiert wird, das heißt es ist

$$\mathcal{G}_n f(x) = \sum_{i=1}^{c(n)} f(x_{i(n)}) \varphi_{i(n)}(x), \quad x \in K, \quad (3.22)$$

mit

$$\varphi_{i(n)}(x_{j(n)}) = \delta_{i(n), j(n)}.$$

Von nun an gehen wir in dieser Arbeit immer von einer Darstellung (3.22) für die exakte kleinste Quadrate Interpolierende aus.

Die angestrebte Konvergenz der Interpolationsoperatoren ist von der Form

$$\lim_{n \rightarrow \infty} \mathcal{G}_n = id_{C(K; \mathbb{R})},$$

genauer $\lim_{n \rightarrow \infty} \mathcal{G}_n f = f$ für alle $f \in C(K; \mathbb{R})$, wobei der Grenzwert $n \rightarrow \infty$ in dem Sinne zu interpretieren ist, dass die Operatorfolge $(\mathcal{G}_n)_{n \geq n_0}$ bezüglich einer monotonen Folge von Interpolationsknotenverteilungen $(\mathcal{B}_n)_{n \geq n_0}$ konstruiert wird.

3.5 Konvergenz bei konstanter monomialer Interpolationsbasis

Dieser Abschnitt untersucht das Konvergenzproblem für exakte kleinste Quadrate Interpolierende, deren Konstruktion auf einer konstanten monomialen Interpolationsbasis der Form $\{b^{(1)}\}$ mit $b^{(1)}(x) = 1$ für alle $x \in K$ basiert. Betrachten wir hierzu die Normalengleichungen (1.39) bzw. deren aufgelöste Form (1.42) wird deutlich, dass in diesem Fall die Interpolierende $\mathcal{G}_n f$ zu vorgelegter Funktion $f \in C(K; \mathbb{R})$, Interpolationsknotenverteilung $\mathcal{B}_n = \{x_1, \dots, x_{c(n)}\}$ und positivem Interpolationsgewicht $w \in C^l(\mathbb{R}^d \setminus \{0\}; \mathbb{R}_{>0})$ durch

$$\mathcal{G}_n f(x) = a_1(x) b^{(1)}(x) = \sum_{i=1}^{c(n)} f(x_i) \frac{w(x - x_i)}{\sum_{j=1}^{c(n)} w(x - x_j)}, \quad x \in K, \quad (3.23)$$

gegeben ist. Das heißt, die Interpolierende $\mathcal{G}_n f$ stimmt mit der Shepard-Interpolierenden $\mathcal{S}_n[w]f$ zu f nach (1.25) überein.

Für stetiges $f \in C(K; \mathbb{R})$ lässt sich die Konvergenz von $\mathcal{G}_n f$ bzw. $\mathcal{S}_n[w]f$ gegen $id_{C(K; \mathbb{R})} f = f$ im Grenzfall $n \rightarrow \infty$ mit Hilfe von Theorem 3.2 zeigen. Der Beweis des folgenden Satzes ist im Wesentlichen eine direkte multivariate Verallgemeinerung von Satz 3.2.

Satz 3.3 *Sei $f \in C(K; \mathbb{R})$ und sei $(\mathcal{B}_n)_{n \geq n_0} \subset K$ eine monotone Folge von Interpolationsknotenverteilungen. Es bezeichne $\mathcal{G}_n f \in C(K; \mathbb{R})$ für $n \geq n_0$ die exakte kleinste Quadrate Interpolierende zu f , gegeben durch*

$$\mathcal{G}_n f(x) = \sum_{i=1}^{c(n)} f(x_i) \frac{w(x - x_i)}{\sum_{j=1}^{c(n)} w(x - x_j)}, \quad x \in K,$$

entsprechend (3.23). Dann gilt

$$\lim_{n \rightarrow \infty} \|\mathcal{G}_n f - f\|_\infty = 0,$$

das heißt, $\mathcal{G}_n f$ konvergiert gleichmäßig gegen f auf K mit $n \rightarrow \infty$.

Zum Beweis von Satz 3.3 benötigen wir folgendes Hilfsresultat:

Lemma 3.5 Für $x = (x(1), \dots, x(d))^T \in \mathbb{R}^d$ bezeichne $e_i^j \in C(\mathbb{R}^d; \mathbb{R})$ die Monomfunktion

$$e_i^j(x) = x(i)^j, \quad x \in \mathbb{R}^d.$$

Dann ist durch

$$\{f_1, \dots, f_{2d+1}\} \equiv \left\{ e_i^j \mid 1 \leq i \leq d, 0 \leq j \leq 2 \right\} \subset C(\mathbb{R}^d; \mathbb{R})$$

eine Menge von positiv summierbaren Funktionen im Sinne der Definition 3.6 gegeben.

Beweis. Für $y = (y(1), \dots, y(d))^T \in \mathbb{R}^d$ setzen wir

$$P_y(x) = \sum_{i=1}^d (x(i) - y(i))^2. \quad (3.24)$$

Dann ist der Ausdruck (3.24) nichtnegativ für alle $x, y \in \mathbb{R}^d$ und kann durch Auflösen der Quadrate in der gesuchten Form

$$P_y(x) = \sum_{i=1}^{2d+1} a_i(y) f_i(x)$$

für gewisse durch direkten Vergleich zu bestimmende Funktionen $a_i \in C(\mathbb{R}^d; \mathbb{R})$ dargestellt werden. Die Eigenschaft $P_y(x) = 0 \iff x = y$ ist sofort einzusehen. Schlussendlich ist

$$\sum_{i=1}^d (2 + 2x(i) + x(i)^2) = \sum_{i=1}^{2d+1} \alpha_i f_i(x)$$

für entsprechende $\alpha_i \in \mathbb{R}$ und für alle $x \in \mathbb{R}^d$ echt positiv. □

Beweis von Satz 3.3 Der Interpolationsoperator \mathcal{G}_n ist im vorliegenden Fall positiv aufgrund der Darstellung (3.23) und der Positivität der unterliegenden Interpolationsgewichte. Daher können wir Theorem 3.2 in dessen Originalversion mit $\Phi_n = \Phi_n^+ = \mathcal{G}_n$ und $\Phi_n^- = 0$ für alle $n \geq n_0$ anwenden. Nach Lemma 3.5 sind die Funktionen e_i^j mit

$$e_i^j : x = (x(1), \dots, x(d))^T \mapsto x(i)^j, \quad 0 \leq j \leq 2, \quad 1 \leq i \leq d, \quad (3.25)$$

positiv summierbar. Können wir

$$\lim_{n \rightarrow \infty} \|\mathcal{G}_n e_i^j - e_i^j\|_\infty = 0, \quad 1 \leq i \leq d, \quad 0 \leq j \leq 2, \quad (3.26)$$

beweisen, folgt mit Hilfe von Theorem 3.2 schon die Existenz von $\lim_{n \rightarrow \infty} \|\mathcal{G}_n f - f\|_\infty = 0$ für beliebiges $f \in C(K; \mathbb{R})$. Der weitere Verlauf des Beweises orientiert sich an dem von Satz 3.2. Hierfür betrachten wir getrennt voneinander die Fälle $j = 0, 1, 2$. Wir verwenden aus Gründen der Übersichtlichkeit die Notation $|x| \equiv \|x\|_2$ für die Euklidische Norm eines Vektors x .

(I) Für $j = 0$ ist nach Konstruktion $\mathcal{G}_n e_i^0 = e_i^0 = 1$ für alle $1 \leq i \leq d$.

(II) Sei nun $\mathcal{B}_n = \{x_1, \dots, x_{c(n)}\}$ und $j = 1$. Dann gelten für $1 \leq i \leq d$ die Ungleichungen

$$\begin{aligned}
& |\mathcal{G}_n e_i^1(x) - e_i^1(x)| \\
& \leq \sum_{j=1}^{c(n)} |x_j(i) - x(i)| \frac{|x - x_j|^{-\alpha}}{\sum_{l=1}^{c(n)} |x - x_l|^{-\alpha}} \\
& \leq \sum_{j=1}^{c(n)} |x_j - x| \frac{|x - x_j|^{-\alpha}}{\sum_{l=1}^{c(n)} |x - x_l|^{-\alpha}} \\
& = \sum_{\substack{j \\ |x - x_j| < r(\mathcal{B}_n)^{d+1}}} |x_j - x| \frac{|x - x_j|^{-\alpha}}{\sum_{l=1}^{c(n)} |x - x_l|^{-\alpha}} \\
& \quad + \sum_{\substack{j \\ |x - x_j| \geq r(\mathcal{B}_n)^{d+1}}} |x_j - x| \frac{|x - x_j|^{-\alpha}}{\sum_{l=1}^{c(n)} |x - x_l|^{-\alpha}} \\
& \leq r(\mathcal{B}_n)^{d+1} + \sum_{\substack{j \\ |x - x_j| \geq r(\mathcal{B}_n)^{d+1}}} |x_j - x| \frac{|x - x_j|^{-\alpha}}{\sum_{l=1}^{c(n)} |x - x_l|^{-\alpha}} \\
& = r(\mathcal{B}_n)^{d+1} + c(n) \mathcal{O}(r(\mathcal{B}_n)^{d+1}) = \mathcal{O}(r(\mathcal{B}_n)^{d+1} + \eta(n)),
\end{aligned}$$

aufgrund von

$$|x - x_j|^{-\alpha+1} = \mathcal{O}(r(\mathcal{B}_n)^{(-\alpha+1)(d+1)})$$

sowie

$$\sum_{l=1}^{c(n)} |x - x_l|^{-\alpha} \geq \sum_{\substack{l \\ |x - x_l| < r(\mathcal{B}_n)^{d+1}}} |x - x_l|^{-\alpha} \geq r(\mathcal{B}_n)^{-\alpha(d+1)}.$$

(III) Schließlich bezeichne $M = \max_{x \in K} |x|$ und sei $j = 2$. Dann gilt für $1 \leq i \leq d$ schon

$$\begin{aligned}
|\mathcal{G}_n e_i^2(x) - e_i^2(x)| & \leq \sum_{j=1}^{c(n)} |(x_j(i))^2 - x(i)^2| \frac{|x - x_j|^{-\alpha}}{\sum_{l=1}^{c(n)} |x - x_l|^{-\alpha}} \\
& \leq 2M \sum_{j=1}^{c(n)} |x_j - x| \frac{|x - x_j|^{-\alpha}}{\sum_{l=1}^{c(n)} |x - x_l|^{-\alpha}} \\
& = \mathcal{O}(r(\mathcal{B}_n)^{d+1} + \eta(n))
\end{aligned}$$

mit Blick auf (II).

Aus (I)-(III) erhalten wir (3.26) durch

$$\|\mathcal{G}_n e_i^j - e_i^j\|_\infty = \sup_{x \in K} |\mathcal{G}_n e_i^j(x) - e_i^j(x)| = \mathcal{O}(r(\mathcal{B}_n)^{d+1} + \eta(n)), \quad 1 \leq i \leq d, \quad 0 \leq j \leq 2,$$

und damit die zu beweisende Konvergenz von $\mathcal{G}_n f$ für beliebiges $f \in C(K; \mathbb{R})$. □

Wird die Ordnung der unterliegenden Interpolationsbasis erhöht, das heißt betrachten wir beispielsweise lineare oder quadratische Basen, geht die Positivität der Operatoren im Allgemeinen verloren. Dies resultiert aus dem Verlust der Positivität der in diesen Fällen konstruierten Kernfunktionen. Theorem 3.2 liefert eine Möglichkeit dieses Problem, unter Verwendung der im folgenden Abschnitt aufgezeigten Eigenschaften der Kernfunktionen, geschickt zu umgehen.

3.6 Eine detailliertere Untersuchung der Kernfunktionen

Zum einen untersuchen wir innerhalb dieses Abschnittes, wie sich die Kernfunktionen beim Übergang zwischen Interpolationsknotenverteilungen der Form \mathcal{B}_n und \mathcal{B}_{n+1} verhalten und gehen zum anderen der Frage nach, wie für festes $x \in K$ der Wert $\varphi_{i(n)}(x)$ für wachsendes n variiert. Diese Untersuchungen werden es uns in Theorem 3.3 erlauben, eine Konvergenzaussage für exakte kleinste Quadrate Interpolierende mit unterliegender quadratischer monomialer Interpolationsbasis mit Hilfe von Theorem 3.2 zu beweisen.

3.6.i Verhalten der Kernfunktionen beim Übergang von \mathcal{B}_n zu \mathcal{B}_{n+1}

Zu $f \in C(K; \mathbb{R})$ sei durch $\mathcal{G}_n f \in C(K; \mathbb{R})$ die exakte kleinste Quadrate Interpolierende bezüglich der Interpolationsbasis $\{b^{(1)}, \dots, b^{(k)}\}$, einem positiven Interpolationsgewicht w und einer Interpolationsknotenverteilung \mathcal{B}_n gegeben. Es bezeichne $\varphi_{i(n)}$ die zugehörige $i(n)$ -te Kernfunktion, das heißt

$$\mathcal{G}_n f = \sum_{i=1}^{c(n)} f(x_{i(n)}) \varphi_{i(n)}$$

mit $\varphi_{i(n)}(x_{j(n)}) = \delta_{i(n),j(n)}$. Wir wollen im Folgenden untersuchen, wie sich die Kernfunktionen beim Übergang von einer Interpolationsknotenverteilung \mathcal{B}_n auf die Verteilung \mathcal{B}_{n+1} verhalten. Ohne Einschränkung betrachten wir den Fall $c(n+1) = c(n) + 1$, das heißt

$$\mathcal{B}_{n+1} = \{x_1, \dots, x_{c(n)}, x_{c(n)+1}\}$$

mit $\mathcal{B}_n = \{x_1, \dots, x_{c(n)}\}$ und $x_{c(n)+1} = x_{c(n+1)}$. Durch $\varphi_{i(n+1)}$ sei die $i(n+1)$ -te Kernfunktion bezüglich \mathcal{B}_{n+1} gegeben, das heißt

$$\varphi_{i(n+1)}(x_{j(n+1)}) = \delta_{i(n),j(n)}$$

für $1 \leq i(n) \leq c(n)$ und $\varphi_{c(n+1)}(x_{j(n+1)}) = \delta_{c(n+1),j(n+1)}$.

Wir gehen der Frage nach, in welchem Zusammenhang die Funktionen $\varphi_{i(n)}$ und $\varphi_{i(n+1)}$ für $1 \leq i(n) \leq c(n)$ stehen. Hierfür betrachten wir zunächst die Shepard-Interpolation als exakte kleinste Quadrate Interpolation mit konstanter monomialer Interpolationsbasis. Nach Satz 3.3 sind die hierbei auftretenden Kernfunktionen für $x \in K$ durch

$$\varphi_{i(n)}(x) = \frac{w(x - x_{i(n)})}{\sum_{j(n)=1}^{c(n)} w(x - x_{j(n)})}, \quad 1 \leq i(n) \leq c(n),$$

sowie für $1 \leq i(n+1) \leq c(n+1) = c(n) + 1$ durch

$$\begin{aligned}\varphi_{i(n+1)}(x) &= \frac{w(x - x_{i(n+1)})}{\sum_{j(n+1)=1}^{c(n+1)} w(x - x_{j(n+1)})} \\ &= \frac{w(x - x_{i(n+1)})}{\sum_{j(n+1)=1}^{c(n)} w(x - x_{j(n+1)}) + w(x - x_{c(n+1)})},\end{aligned}\quad (3.27)$$

gegeben. Für diesen Fall gilt offensichtlich

$$\varphi_{i(n+1)}(x) \leq \varphi_{i(n)}(x), \quad 1 \leq i(n) = i(n+1) \leq c(n). \quad (3.28)$$

Durch elementaren Umformungen erhalten wir

$$\varphi_{i(n)}(x) - \varphi_{i(n+1)}(x) = \frac{w(x - x_{i(n)}) \left(\sum_{j(n)=1}^{c(n)} w(x - x_{j(n)}) \right)^{-2}}{\left(w(x - x_{c(n+1)}) \right)^{-1} + \left(\sum_{j(n)=1}^{c(n)} w(x - x_{j(n)}) \right)^{-1}}. \quad (3.29)$$

Obwohl die Gültigkeit der Ungleichung (3.28) sofort durch einen Vergleich von $\varphi_{i(n)}$ mit $\varphi_{i(n+1)}$ folgt, erhalten wir diese auch durch (3.29), denn die rechte Seite der Gleichung (3.29) ist immer positiv.

Eine Untersuchung der Kernfunktionen für den allgemeinen Fall liefert ein vergleichbares Ergebnis als direkte Verallgemeinerung der soeben aufgezeigten Relation (3.29):

In Abschnitt 1.6 hatten wir gesehen, dass für $x \in K$ die Darstellung

$$\varphi_{i(n)}(x) = w(x - x_{i(n)}) \underline{b}(x_{i(n)})^T S(x) \underline{b}(x) \quad (3.30)$$

mit $S(x) = (BW(x)B^T)^{-1} \in \mathbb{R}^{k \times k}$ Gültigkeit besitzt. Hierbei ist für $x \in \mathcal{B}_n$ die Matrix $S(x)$ im Sinne von Lemma 1.4 zu bestimmen und es gilt $\varphi_{i(n)}(x_{i(n)}) = 1$ nach Satz 1.10. Die Kernfunktion $\varphi_{i(n+1)}(x)$ besitzt in Analogie zu (3.30) die Darstellung

$$\varphi_{i(n+1)}(x) = w(x - x_{i(n+1)}) \underline{b}(x_{i(n+1)})^T \left(\widetilde{BW(x)B^T} \right)^{-1} \underline{b}(x)$$

mit

$$\widetilde{BW(x)B^T}$$

$$= \begin{pmatrix} \sum_{i(n+1)=1}^{c(n+1)} b^{(1)}(x_i) w(x - x_i) b^{(1)}(x_i) & \dots & \sum_{i(n+1)=1}^{c(n+1)} b^{(1)}(x_i) w(x - x_i) b^{(k)}(x_i) \\ \vdots & \ddots & \vdots \\ \sum_{i(n+1)=1}^{c(n+1)} b^{(k)}(x_i) w(x - x_i) b^{(1)}(x_i) & \dots & \sum_{i(n+1)=1}^{c(n+1)} b^{(k)}(x_i) w(x - x_i) b^{(k)}(x_i) \end{pmatrix}$$

$$= BW(x)B^T + w(x - x_{c(n+1)}) \underline{b}(x_{c(n+1)}) \underline{b}(x_{c(n+1)})^T,$$

das heißt wir erhalten unter Zuhilfenahme der Sherman-Morrison-Formel (vgl. Lemma 1.3) für $1 \leq i(n+1) \leq c(n+1)$ die Identität

$$\begin{aligned} \varphi_{i(n+1)}(x) &= w(x - x_{i(n+1)}) \underline{b}(x_{i(n+1)})^T \left(BW(x) B^T + w(x - x_{c(n+1)}) \underline{b}(x_{c(n+1)}) \underline{b}(x_{c(n+1)})^T \right)^{-1} \underline{b}(x) \\ &= w(x - x_{i(n+1)}) \underline{b}(x_{i(n+1)})^T \left(S(x) - \frac{w(x - x_{c(n+1)}) \left(S(x) \underline{b}(x_{c(n+1)}) \right) \left(S(x) \underline{b}(x_{c(n+1)}) \right)^T}{1 + w(x - x_{c(n+1)}) \underline{b}(x_{c(n+1)})^T S(x) \underline{b}(x_{c(n+1)})} \right)^T \underline{b}(x) \end{aligned}$$

mit $S(x) = \left(BW(x) B^T \right)^{-1}$. Dies liefert für $1 \leq i(n+1) = i(n) \leq c(n)$

$$\varphi_{i(n+1)}(x) = \varphi_{i(n)}(x) - \frac{w(x - x_{i(n)}) \left(\underline{b}(x_{i(n)})^T S(x) \underline{b}(x_{c(n+1)}) \right) \left(\underline{b}(x_{c(n+1)})^T S(x) \underline{b}(x) \right)}{\left(w(x - x_{c(n+1)}) \right)^{-1} + \underline{b}(x_{c(n+1)})^T S(x) \underline{b}(x_{c(n+1)})}. \quad (3.31)$$

Die Gleichung (3.29) ist ein Spezialfall von (3.31), denn für den Fall der multivariaten Shepard-Interpolation ist $\underline{b}(x) = 1$ für alle $x \in K$ und es gilt

$$S(x) = \left(\sum_{i(n)=1}^{c(n)} w(x - x_{i(n)}) \right)^{-1}.$$

3.6.ii Asymptotik der Kernfunktionen für $n \rightarrow \infty$

Im Folgenden gehen wir von analogen Voraussetzungen an die Kernfunktionen $\varphi_{i(n)}$ wie in Abschnitt 3.6.i aus. In Satz 1.10 hatten wir bewiesen, dass die $i(n)$ -te Kernfunktion mit der Darstellung

$$\varphi_{i(n)}(x) = w(x - x_{i(n)}) \underline{b}(x_{i(n)})^T \left(BW(x) B^T \right)^{-1} \underline{b}(x), \quad x \in K, \quad (3.32)$$

einer Kronecker-Delta-Eigenschaft genügt. Das heißt für jedes $x_{j(n)} \in \mathcal{B}_n$ gilt

$$\lim_{x \rightarrow x_{j(n)}} \varphi_{i(n)}(x) = \delta_{i(n), j(n)}. \quad (3.33)$$

Die Eigenschaft (3.33) garantiert uns zusammen mit Satz 3.4 ein punktweises Verschwinden der einzelnen Kernfunktionen $\varphi_{i(n)}(x)$ für jedes $x \neq x_{i(n)}$ im Grenzwert $n \rightarrow \infty$.

Definition 3.8 Für eine Matrix $M \in \mathbb{R}^{p \times q}$ mit $M = (m_{i,j})_{\substack{i=1..p \\ j=1..q}}$ bezeichnet $M|_{\{i,j\}}$ das Matrix-Element $m_{i,j}$.

Satz 3.4 Sei $\{b^{(1)}, \dots, b^{(k)}\} \subset C^m(K; \mathbb{R})$ eine Interpolationsbasis und $w \in C^l(\mathbb{R}^d \setminus \{0\}; \mathbb{R}_{>0})$ ein (Norm-inverses) positives Interpolationsgewicht von der Form $w(x) = \|x\|_2^{-\alpha}$. Weiter sei für eine Interpolationsknotenverteilung $\mathcal{B}_n = \{x_1, \dots, x_{c(n)}\} \subset K$ die Matrix $B \in \mathbb{R}^{k \times c(n)}$ durch $B|_{\{i,j\}} = b^{(i)}(x_j)$ für $1 \leq i \leq k$ und $1 \leq j \leq c(n)$ sowie die Matrix $W(x) \in \mathbb{R}^{c(n) \times c(n)}$ vermöge $W(x) = \text{diag}(w(x - x_1), \dots, w(x - x_{c(n)}))$ definiert.

Dann gilt für beliebiges $x \in K \setminus \mathcal{B}_n$ und $1 \leq k_1, k_2 \leq k$

$$\left(BW(x)B^T \right)^{-1} \Big|_{\{k_1, k_2\}} = \mathcal{O}(c(n)^{k-1} r(\mathcal{B}_n)^\alpha).$$

Beweis.

Untersuchen wir zunächst den Fall $k = 1$. Es sei $x_{i_0} \in \mathcal{B}_n$ mit $\|x - x_{i_0}\|_2 < r(\mathcal{B}_n)$. Dann ist $BW(x)B^T = \sum_{j=1}^{c(n)} w(x - x_j)$ und die Behauptung folgt mit

$$\frac{1}{\sum_{j=1}^{c(n)} w(x - x_j)} \leq \frac{1}{w(x - x_{i_0})} = \|x - x_{i_0}\|_2^\alpha = \mathcal{O}(r(\mathcal{B}_n)^\alpha).$$

Ist $k > 1$ nutzen wir die Adjunkten-Regel zur Invertierung der Matrix $BW(x)B^T$. Aufgrund von Lemma 1.4 wissen wir, dass $(BW(x)B^T)^{-1} \Big|_{\{k_1, k_2\}} \in \mathbb{R}$ für jedes $x \in K$ ist. Setzen wir $S(x) = BW(x)B^T$, berechnet sich die Inverse $S(x)^{-1}$ durch

$$S(x)^{-1} = \frac{1}{\det S(x)} \text{Adj} S(x),$$

wobei $\text{Adj} S(x)$ die Adjunkte der Matrix $S(x)$ bezeichnet. Jeder Eintrag $\text{Adj} S(x) \Big|_{\{k_1, k_2\}}$ der Adjunkten ist von der Form

$$\text{Adj} S(x) \Big|_{\{k_1, k_2\}} = \sum_{\tilde{j}_1=1}^{c(n)} \cdots \sum_{\tilde{j}_{k-1}=1}^{c(n)} \tilde{b}_{\tilde{j}_1, \dots, \tilde{j}_{k-1}}^{(k_1, k_2)} w(x - x_{\tilde{j}_1}) \cdots w(x - x_{\tilde{j}_{k-1}})$$

mit $\tilde{b}_{\tilde{j}_1, \dots, \tilde{j}_{k-1}}^{(k_1, k_2)} \in \mathbb{R}$ und elementare Umformungen zeigen

$$\begin{aligned} \det S(x) &= \det \begin{pmatrix} \sum_{j=1}^{c(n)} b^{(1)}(x_j) w(x - x_j) b^{(1)}(x_j) & \cdots & \sum_{j=1}^{c(n)} b^{(1)}(x_j) w(x - x_j) b^{(k)}(x_j) \\ \vdots & \ddots & \vdots \\ \sum_{j=1}^{c(n)} b^{(k)}(x_j) w(x - x_j) b^{(1)}(x_j) & \cdots & \sum_{j=1}^{c(n)} b^{(k)}(x_j) w(x - x_j) b^{(k)}(x_j) \end{pmatrix} \\ &= \sum_{j_1=1}^{c(n)} \cdots \sum_{j_k=1}^{c(n)} w(x - x_{j_1}) \cdots w(x - x_{j_k}) b^{(1)}(x_{j_1}) \cdots b^{(k)}(x_{j_k}) \det(\underline{b}(x_{j_1}), \dots, \underline{b}(x_{j_k})) \\ &= \frac{1}{k!} \sum_{j_1=1}^{c(n)} \cdots \sum_{j_k=1}^{c(n)} w(x - x_{j_1}) \cdots w(x - x_{j_k}) (\det(\underline{b}(x_{j_1}), \dots, \underline{b}(x_{j_k})))^2. \end{aligned}$$

Ist nun $x_{i_1}, x_{i_2}, \dots, x_{i_{c(n)}}$ eine Ordnung der Interpolationsknoten mit

$$w(x - x_{i_1}) \geq w(x - x_{i_2}) \geq \cdots \geq w(x - x_{i_{c(n)}}),$$

so gilt

$$\text{Adj} S(x) \Big|_{\{k_1, k_2\}} = \mathcal{O}(c(n)^{k-1} w(x - x_{i_2}) \cdots w(x - x_{i_k}))$$

und

$$\det S(x) \geq \frac{1}{k!} w(x - x_{i_1}) \cdots w(x - x_{i_k}) (\det(\underline{b}(x_{i_1}), \dots, \underline{b}(x_{i_k})))^2,$$

das heißt für jeden Eintrag der Inversen $S(x)^{-1}$ ist

$$S(x)^{-1}|_{\{k_1, k_2\}} = \mathcal{O}(c(n)^{k-1} r(\mathcal{B}_n)^\alpha).$$

□

Anhand von Satz 3.4 erkennen wir sofort, dass aufgrund von (3.32) für jedes $x \in K \setminus \mathcal{B}_n$ die Relation

$$\varphi_{i(n)}(x) = \mathcal{O}(c(n)^{k-1} r(\mathcal{B}_n)^\alpha) \quad (3.34)$$

gilt. Mit Hilfe von Theorem 2.1 kann eine Aussage über die Lipschitz-Konstanten der Kernfunktionen für den Fall $\min\{m, l\} \geq 1$ getroffen werden, wodurch das Resultat (3.34) qualitativ zu verbessern ist.

Satz 3.5 *Unter den Voraussetzungen von Satz 3.4 und $\min\{m, l\} \geq 1$ sei durch*

$$\varphi_{i(n)}(x) = w(x - x_{i(n)}) \underline{b}(x_{i(n)})^T \left(BW(x) B^T \right)^{-1} \underline{b}(x), \quad x \in K, \quad n \geq n_0,$$

die $i(n)$ -te Kernfunktion gegeben. Dann erfüllt $\varphi_{i(n)}$ auf $K \setminus B_{r(\mathcal{B}_n)}(x_{i(n)})$ eine lokale Lipschitz-Bedingung

$$|\varphi_{i(n)}(\xi_1) - \varphi_{i(n)}(\xi_2)| \leq L_{\varphi_{i(n)}} \|\xi_1 - \xi_2\|_2, \quad \xi_1, \xi_2 \in K \setminus B_{r(\mathcal{B}_n)}(x_{i(n)}),$$

mit einer Lipschitz-Konstanten $L_{\varphi_{i(n)}} = \mathcal{O}(c(n)^{k-1} r(\mathcal{B}_n)^\alpha)$.

Beweis. Aufgrund der Gleichungen (2.4) und (2.5) in Theorem 2.1 gilt für jeden Multi-Index μ von der Ordnung $|\mu| = 1$ schon

$$D^\mu \varphi_{i(n)}(x) = \mathcal{O}(c(n)^{k-1} r(\mathcal{B}_n)^\alpha)$$

für $x \in K \setminus B_{r(\mathcal{B}_n)}(x_{i(n)})$ unter direkter Beachtung von Satz 3.4. Dann gilt für die Lipschitz-Konstante die Relation

$$L_{\varphi_{i(n)}} \leq \sum_{k=1}^d \sup_{\xi \in K \setminus B_{r(\mathcal{B}_n)}(x_{i(n)})} |D^{\mu_k} \varphi_{i(n)}(\xi)|$$

für die Multi-Indizes $\mu_k = (\delta_{1,k}, \delta_{2,k}, \dots, \delta_{d-1,k}, \delta_{d,k})$, das heißt $L_{\varphi_{i(n)}} = \mathcal{O}(c(n)^{k-1} r(\mathcal{B}_n)^\alpha)$. □

Jetzt sind sämtliche Hilfsmittel bereitgestellt, um das punktweise Verschwinden der einzelnen Kernfunktionen $\varphi_{i(n)}(x)$ für jedes $x \neq x_{i(n)}$ im Grenzfall $n \rightarrow \infty$ zu beweisen.

Satz 3.6 *Seien die $\varphi_{i(n)} \in C^{\min\{m, l\}}(K; \mathbb{R})$ für $n \geq n_0$ Kernfunktionen konstruiert im Sinne der Definition 1.9 zu vorgelegter Interpolationsbasis $\{b^{(1)}, \dots, b^{(k)}\} \subset C^m(K; \mathbb{R})$, einem positiven Interpolationsgewicht $w \in C^l(\mathbb{R}^d \setminus \{0\}; \mathbb{R}_{>0})$ der Form $w(x) = \|x\|_2^{-\alpha}$ sowie einer unterliegenden monotonen Folge von Interpolationsknotenverteilungen $(\mathcal{B}_n)_{n \geq n_0}$ in K . Dann gilt für $x \neq x_{i(n)}$*

$$\lim_{n \rightarrow \infty} |\varphi_{i(n)}(x)| = 0,$$

im Fall

- (i) $k = 1$,
- (ii) $k = 2$, falls $\alpha \geq d$,
- (iii) $k \geq 3$, falls die Folge $(c(n)^{k-2}r(\mathcal{B}_n)^{\alpha-d})_{n \geq n_0}$ beschränkt ist.

Beweis. Zunächst wählen wir ein $N_0 \in \mathbb{N}$ so groß, dass $x \in K \setminus B_{r(\mathcal{B}_n)}(x_{i(n)})$ für $n \geq N_0$ ist. Im weiteren Verlauf sei $n \geq N_0$ angenommen. Dann existiert ein $x_{j(n)} \in \mathcal{B}_n$ mit $x_{j(n)} \in B_{r(\mathcal{B}_n)}(x)$ und $x_{j(n)} \neq x_{i(n)}$. Aufgrund der Kronecker-Delta-Eigenschaft

$$\lim_{x \rightarrow x_{j(n)}} \varphi_{i(n)}(x) = \delta_{i(n), j(n)}$$

der Kernfunktionen gilt

$$|\varphi_{i(n)}(x)| = |\varphi_{i(n)}(x) - \varphi_{i(n)}(x_{j(n)})| \leq L_{\varphi_{i(n)}} \|x - x_{j(n)}\|_2 \leq L_{\varphi_{i(n)}} r(\mathcal{B}_n). \quad (3.35)$$

Nach Satz 3.5 ist $L_{\varphi_{i(n)}} = \mathcal{O}(c(n)^{k-1}r(\mathcal{B}_n)^\alpha)$, das heißt mit Gleichung (3.35) erhalten wir

$$|\varphi_{i(n)}(x)| = \mathcal{O}(c(n)^{k-1}r(\mathcal{B}_n)^{\alpha+1}).$$

Ist nun $k = 1$ folgt die Behauptung aus $\lim_{n \rightarrow \infty} r(\mathcal{B}_n) = 0$ und für $k \geq 2$ gilt

$$c(n)^{k-1}r(\mathcal{B}_n)^{\alpha+1} = c(n)r(\mathcal{B}_n)^{d+1}c(n)^{k-2}r(\mathcal{B}_n)^{\alpha-d},$$

was mit Blick auf die Definition der Folge $(\mathcal{B}_n)_{n \geq n_0}$ ebenfalls die Behauptung liefert. \square

3.7 Konvergenz bei quadratischer monomialer Interpolationsbasis

Die folgenden Betrachtungen sind der Untersuchung des Konvergenzproblems für die exakte kleinste Quadrate Interpolation mit unterliegender multivariater quadratischer Interpolationsbasis der Form

$$\{b^{(1)}, \dots, b^{(k)}\} = \{e_i^j \mid 1 \leq i \leq d, 0 \leq j \leq 2\} \quad (3.36)$$

mit den Monomfunktionen e_i^j , wie in Gleichung (3.25) eingeführt, gewidmet. Für die Interpolationsbasisfunktionen gilt hierbei insbesondere $b^{(i)} \in C^\infty(K; \mathbb{R})$, das heißt es ist $m = \infty$.

Für jede der Kernfunktionen $\varphi_{i(n)}$ existiert als stetige Funktion eine natürliche Darstellung

$$\varphi_{i(n)} = \varphi_{i(n)}^\oplus - \varphi_{i(n)}^\ominus, \quad 1 \leq i \leq c(n), \quad (3.37)$$

wobei die Funktionen $\varphi_{i(n)}^\oplus$ und $\varphi_{i(n)}^\ominus$ punktweise für $x \in K$ durch

$$\varphi_{i(n)}^\oplus(x) = \max\{\varphi_{i(n)}(x), 0\},$$

$$\varphi_{i(n)}^\ominus(x) = \max\{-\varphi_{i(n)}(x), 0\}$$

definiert sind. Offensichtlich sind die Funktionen $\varphi_{i(n)}^\oplus$ und $\varphi_{i(n)}^\ominus$ sowohl nichtnegativ als auch stetig.

Aufgrund der Identität (3.37) kann der Interpolationsoperator \mathcal{G}_n durch

$$\begin{aligned}\mathcal{G}_n f &= \sum_{i=1}^{c(n)} f(x_{i(n)}) \varphi_{i(n)} = \sum_{i=1}^{c(n)} f(x_{i(n)}) \varphi_{i(n)}^{\oplus} - \sum_{i=1}^{c(n)} f(x_{i(n)}) \varphi_{i(n)}^{\ominus} \\ &=: \mathcal{G}_n^+ f - \mathcal{G}_n^- f\end{aligned}$$

für beliebiges $f \in C(K; \mathbb{R})$ dargestellt werden und es resultiert für $\mathcal{G}_n : C(K; \mathbb{R}) \rightarrow C(K; \mathbb{R})$ die Darstellung

$$\mathcal{G}_n = \mathcal{G}_n^+ - \mathcal{G}_n^- \quad (3.38)$$

für positive lineare Operatoren $\mathcal{G}_n^+, \mathcal{G}_n^- : C(K; \mathbb{R}) \rightarrow C(K; \mathbb{R})$.

Aufgrund der Darstellung (3.37) leiten wir aus Satz 3.6 das folgende Korollar ab.

Korollar 3.1 Für die Kernfunktionen $\varphi_{i(n)} \in C^{\min\{m,l\}}(K; \mathbb{R})$ aus Satz 3.6 gilt

$$\varphi_{i(n)}^{\ominus}(x) = \mathcal{O}(c(n)^{k-1} r(\mathcal{B}_n)^{\alpha+1}), \quad x \in K. \quad (3.39)$$

Beweis. Die Behauptung folgt aus (3.37) sowie dem Beweis zu Satz 3.6. Insbesondere ist jede Kernfunktion $\varphi_{i(n)}$ in einer genügend kleinen Umgebung von $x_{i(n)}$ echt positiv. \square

Setzen wir für die Kernfunktionen lediglich $\varphi_{i(n)} \in C(K; \mathbb{R})$ voraus, so gilt entsprechend (3.34) anstatt von (3.39) die Relation

$$\varphi_{i(n)}^{\ominus}(x) = \mathcal{O}(c(n)^{k-1} r(\mathcal{B}_n)^{\alpha}), \quad x \in K.$$

Theorem 3.3 Sei $f \in C(K; \mathbb{R})$ und $(\mathcal{B}_n)_{n \geq n_0}$ eine monotone Folge von Interpolationsknotenverteilungen in K . Es bezeichne $\mathcal{G}_n f \in C(K; \mathbb{R})$ die exakte kleinste Quadrate Interpolierende zu f , die mit Hilfe der multivariaten Interpolationsbasis (3.36) und einem positiven Interpolationsgewicht $w \in C^l(\mathbb{R}^d \setminus \{0\}; \mathbb{R}_{>0})$ mit $w(x) = \|x\|_2^{-\alpha}$ konstruiert wird. Falls $(c(n)^{2d} r(\mathcal{B}_n)^{\alpha-d})_{n \geq n_0}$ eine beschränkte Folge ist, dann gilt

$$\lim_{n \rightarrow \infty} \|\mathcal{G}_n f - f\|_{\infty} = 0,$$

das heißt $\mathcal{G}_n f$ konvergiert gleichmäßig gegen f auf K für $n \rightarrow \infty$.

Beweis. Nach Lemma 3.5 ist durch die Menge der Interpolationsbasisfunktionen

$$\left\{ e_i^j \mid 1 \leq i \leq d, 0 \leq j \leq 2 \right\}$$

eine Menge von positiv summierbaren Funktionen gegeben. Ferner wird jede Basisfunktion e_i^j vom Interpolationsoperator \mathcal{G}_n exakt reproduziert, wie wir in Satz 1.11 bewiesen haben. Für beliebiges $n \geq n_0$ ist

$$\mathcal{G}_n f = \sum_{i=1}^{c(n)} f(x_{i(n)}) \varphi_{i(n)} = \sum_{i=1}^{c(n)} f(x_{i(n)}) \varphi_{i(n)}^{\oplus} - \sum_{i=1}^{c(n)} f(x_{i(n)}) \varphi_{i(n)}^{\ominus}$$

mit

$$\sum_{i=1}^{c(n)} f(x_{i(n)}) \varphi_{i(n)}^{\ominus}(x) = c(n) \mathcal{O}(c(n)^{(2d+1)-1} r(\mathcal{B}_n)^{\alpha+1}) = \mathcal{O}(\eta(n) c(n)^{2d} r(\mathcal{B}_n)^{\alpha-d})$$

für alle $x \in K$ nach Korollar 3.1, das heißt es gilt

$$\lim_{n \rightarrow \infty} \sum_{i=1}^{c(n)} f(x_{i(n)}) \varphi_{i(n)}^{\ominus}(x) = 0$$

für beliebiges $x \in K$ aufgrund der Voraussetzung an die Folge $(c(n)^{2d} r(\mathcal{B}_n)^{\alpha-d})_{n \geq n_0}$. In der Konsequenz ist durch $(\mathcal{G}_n)_{n \geq n_0}$ eine Folge von fast positiven linearen Operatoren gegeben und Theorem 3.2 liefert das zu beweisende Konvergenzresultat. \square

Theorem 3.3 liefert eine reine Konvergenzaussage für exakte kleinste Quadrate Interpolierende. Jedoch macht es über die Güte der Approximation von $\mathcal{G}_n f$ an f in Abhängigkeit von $n \geq n_0$ keine Aussage. Wir bringen nun einen alternativen Beweis für das Theorem 3.3, der auf einer derartigen Gütebetrachtung durch einen Stetigkeitsmodul beruht.

Definition 3.9 Für $f \in C(K; \mathbb{R})$ und $x \in K$ bezeichnet

$$\Delta_t f(x) = f(x+t) - f(x)$$

für $t \in \mathbb{R}^d$ mit $x+t \in K$ die 1-te Differenz von f bei x in Richtung t . Dann ist der Stetigkeitsmodul $\omega : C(K; \mathbb{R}) \times \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}_{\geq 0}$ durch

$$\omega(f, h) = \sup_{0 \leq |t| \leq h} |\Delta_t f| = \sup_{0 \leq |t| \leq h} \sup_{x \in K} |\Delta_t f(x)| \quad (3.40)$$

gegeben, wobei h , das Argument des Stetigkeitsmoduls, maximal so groß zu wählen ist, dass die Differenz $\Delta_t f(x)$ für jedes $x \in K$ ausgewertet werden kann.

Das multivariate Äquivalent zu Lemma 3.3(5) liefert Lemma 3.6, dessen Beweis in der Monographie [3] zu finden ist.

Lemma 3.6 Sei $f \in C(K; \mathbb{R})$ und ω das Stetigkeitsmodul wie in (3.40). Dann gilt

$$\omega(f, st) \leq (1 + [s]) \omega(f, t),$$

für beliebige $s, t \in \mathbb{R}_{>0}$. Hierbei bezeichnet $[\cdot]$ die Gauß-Klammer, das heißt

$$[s] = \max\{n \in \mathbb{N} \mid n \leq s\}, \quad s \in \mathbb{R}_{>0}.$$

In Abschnitt 3.1 hatten wir bereits eine Ungleichung vom Cauchy-Schwarz-Typ für positive und lineare Operatoren erwähnt. Genauer gilt für diese Operatoren $P : C(K; \mathbb{R}) \rightarrow C(K; \mathbb{R})$ die Ungleichung

$$P(fg) \leq \sqrt{P(f^2)} \sqrt{P(g^2)}, \quad f, g \in C(K; \mathbb{R}), \quad (3.41)$$

die punktweise für $x \in K$ zu interpretieren ist. Der Beweis dieser Ungleichung kommt recht klassisch daher: Für $\lambda \in \mathbb{R}$ ist $(f + \lambda g)^2 \geq 0$, was aufgrund der Positivität des Operators und dessen Linearität $P(f^2) + 2\lambda P(fg) + \lambda^2 P(g^2) \geq 0$ impliziert. Eine genauere Untersuchung der quadratischen Ungleichung in λ liefert das Ergebnis.

Theorem 3.4 Sei $f \in C(K; \mathbb{R})$ und $\mathcal{G}_n f \in C(K; \mathbb{R})$ für $n \geq n_0$ die exakte kleinste Quadrate Interpolierende wie in Theorem 3.3. Zu $x, y \in \mathbb{R}^d$ sei

$$P_y(x) = \|x - y\|_2^2 = \sum_{i=1}^d (x(i) - y(i))^2$$

und für $p(y)(\cdot) = P_y(\cdot)$ sei

$$\alpha_n(x) = \sqrt{\mathcal{G}_n^-[p(y)](x)}, \quad x \in K.$$

Dann gilt für beliebiges $x \in K$ die Ungleichung

$$|\mathcal{G}_n f(x) - f(x)| \leq |\mathcal{G}_n^-[p(y)](x)| + 2 \cdot \omega(f, \alpha_n(x)). \quad (3.42)$$

Beweis. Sei $\Phi : C(K; \mathbb{R}) \rightarrow C(K; \mathbb{R})$ ein positiver linearer Operator. Für diesen gilt

$$\left| \Phi[f](x) - f(x) \right| \leq \underbrace{\left| \Phi[f](x) - \Phi[f(x) \cdot f_1(y)](x) \right|}_{=:A} + \underbrace{\left| \Phi[f(x) \cdot f_1(y)](x) - f(x) \cdot f_1(x) \right|}_{=:B}.$$

Eine Analyse des B-Terms ist recht trivial, denn

$$B = |f(x)| \cdot \left| \Phi[f_1(y)](x) - f_1(x) \right|.$$

Eine Abschätzung von A ergibt sich für $t > 0$ durch

$$\begin{aligned} A &= \left| \Phi[f(y) - f(x) \cdot f_1(y)](x) \right| \leq \Phi[|f(y) - f(x) \cdot f_1(y)|](x) \\ &\leq \Phi\left[\left(1 + \frac{1}{t} \|y - x\|_2\right)\right](x) \cdot \omega(f, t) \\ &= \left(\Phi[f_1(y)](x) + \frac{1}{t} \Phi[\|y - x\|_2](x) \right) \cdot \omega(f, t), \end{aligned}$$

wobei zum einen Lemma 3.6, das heißt

$$\left| f(y) - f(x) \cdot f_1(y) \right| \leq \omega(f, \|y - x\|_2) \leq \left(1 + \frac{1}{t} \|y - x\|_2\right) \cdot \omega(f, t),$$

und zum anderen die Monotonie von Φ Anwendung findet. Weiterhin liefert die Cauchy-Schwarz-Ungleichung (3.41) für positive lineare Operatoren

$$\Phi[\|y - x\|_2](x) \leq \sqrt{\Phi[p(y)](x)} \cdot \sqrt{\Phi[f_1(y)](x)},$$

und es folgt

$$\begin{aligned} \left| \Phi[f](x) - f(x) \right| &\leq \left(\Phi[f_1(y)](x) + \frac{\sqrt{\Phi[p(y)](x)}}{t} \sqrt{\Phi[f_1(y)](x)} \right) \cdot \omega(f, t) \\ &\quad + |f(x)| \cdot \left| \Phi[f_1(y)](x) - f_1(x) \right|. \end{aligned} \quad (3.43)$$

Setzen wir jetzt $\Phi = \mathcal{G}_n^+$ und $t = \alpha_n(x)$ in der Ungleichung (3.43) und verwenden die Identität $\mathcal{G}_n = \mathcal{G}_n^+ - \mathcal{G}_n^-$ beziehungsweise $\mathcal{G}_n^+ = \mathcal{G}_n + \mathcal{G}_n^-$, liefert dies

$$\begin{aligned} & \left| (\mathcal{G}_n + \mathcal{G}_n^-)f(x) - f(x) \right| \\ & \leq \left((\mathcal{G}_n + \mathcal{G}_n^-)f_1(x) + \frac{\sqrt{(\mathcal{G}_n + \mathcal{G}_n^-)P_y(x)}}{\alpha_n(x)} \sqrt{(\mathcal{G}_n + \mathcal{G}_n^-)f_1(x)} \right) \cdot \omega(f, \alpha_n(x)) \\ & \quad + |f(x)| \cdot \left| (\mathcal{G}_n + \mathcal{G}_n^-)f_1(x) - f_1(x) \right| \\ & = \left(f_1(x) + \mathcal{G}_n^- f_1(x) + \frac{\sqrt{\mathcal{G}_n^- P_y(x)}}{\alpha_n(x)} \sqrt{f_1(x) + \mathcal{G}_n^- f_1(x)} \right) \cdot \omega(f, \alpha_n(x)) \\ & \quad + |f(x)| \cdot \left| \mathcal{G}_n^- f_1(x) \right| = 2 \cdot \omega(f, \alpha_n(x)). \end{aligned}$$

Die vorausgehenden Umformungen sind gültig aufgrund von $\mathcal{G}_n f_1(x) = f_1(x)$, das heißt insbesondere $\mathcal{G}_n^- f_1(x) = 0$ und

$$\mathcal{G}_n P_y(x) = \mathcal{G}_n[p(y)](x) = 0.$$

Diese letzte Eigenschaft folgt aus der exakten Reproduktion von $P_y(x)$ durch \mathcal{G}_n und $P_x(x) = 0$. Nun folgt (3.42) aus den obigen Betrachtungen sowie einer Anwendung der umgekehrten Dreiecks-Ungleichung auf den Ausdruck

$$\left| (\mathcal{G}_n + \mathcal{G}_n^-)f(x) - f(x) \right|.$$

□

Korollar 3.2 *Unter den Voraussetzungen von Theorem 3.4 gilt*

$$\|\mathcal{G}_n f - f\|_\infty \leq G_n + 2 \cdot W_n$$

mit $G_n = \sup_{x \in K} |\mathcal{G}_n^- f(x)|$ und $W_n = \sup_{x \in K} |\omega(f, \alpha_n(x))|$ für $n \geq n_0$.

3.8 Konvergenz bei linearer monomialer Interpolationsbasis

Zum Abschluss des dritten Kapitels dieser Arbeit wollen wir der Konvergenzfrage für exakte kleinste Quadrate Interpolierende für eine unterliegende univariate lineare monomiale Interpolationsbasis nachgehen. Der Beweis des hierfür gültigen Resultates orientiert sich mit leichten Modifikationen an den Ideen von Theorem 3.1.

Zunächst geben wir jedoch analog zu (3.23) eine explizite Darstellung dieser exakte kleinste Quadrate Interpolierenden an.

Satz 3.7 *Sei $f \in C([a, b]; \mathbb{R})$ und $\{x_1, \dots, x_n\} \subset [a, b]$ eine Interpolationsknotenverteilung. Dann ist die exakte kleinste Quadrate Interpolierende zu f bezüglich der Interpolationsbasis $\{e_0, e_1\}$ mit $e_k(x) = x^k$ und einem positiven Interpolationsgewicht $w \in C^l(\mathbb{R} \setminus \{0\}; \mathbb{R}_{>0})$ durch*

$$\mathcal{G}f(x) = \sum_{j=1}^n f(x_j) \varphi_j(x), \quad x \in K,$$

mit

$$\varphi_j(x) = \frac{\sum_{i=1}^n w(x-x_j)w(x-x_i)(x_i^2 + x_j(x-x_i) - xx_i)}{\sum_{i=1}^n \sum_{k=1}^n w(x-x_i)w(x-x_k)(x_i^2 - x_ix_k)}$$

gegeben.

Beweis. Nach Gleichung (1.62) gilt

$$\varphi_i(x) = \lim_{\varepsilon \rightarrow 0} w^{(i)}(x + \varepsilon) \underline{b}(x_i)^T \left(BW(x + \varepsilon) B^T \right)^{-1} \underline{b}(x)$$

für die Matrix

$$B = \begin{pmatrix} 1 & 1 & \dots & 1 \\ x_1 & x_2 & \dots & x_n \end{pmatrix},$$

die Diagonalmatrix

$$W(\bar{x}) = \text{diag}(w(\bar{x} - x_1), \dots, w(\bar{x} - x_n)),$$

und dem Vektor $\underline{b}(x) = (1, x)^T$. Aufgrund von

$$W(x)B^T = \begin{pmatrix} w(x-x_1) & w(x-x_1)x_1 \\ \vdots & \vdots \\ w(x-x_n) & w(x-x_n)x_n \end{pmatrix},$$

$$BW(x)B^T = \begin{pmatrix} \sum_{i=1}^n w(x-x_i) & \sum_{i=1}^n w(x-x_i)x_i \\ \sum_{i=1}^n w(x-x_i)x_i & \sum_{i=1}^n w(x-x_i)x_i^2 \end{pmatrix}$$

und

$$\left(BW(x)B^T \right)^{-1} = \frac{1}{\det(x)} \begin{pmatrix} \sum_{i=1}^n w(x-x_i)x_i^2 & -\sum_{i=1}^n w(x-x_i)x_i \\ -\sum_{i=1}^n w(x-x_i)x_i & \sum_{i=1}^n w(x-x_i) \end{pmatrix}$$

mit $\det(x) = \sum_{i=1}^n \sum_{k=1}^n w(x-x_i)w(x-x_k)(x_i^2 - x_ix_k)$ folgt die Behauptung. \square

Eine Konvergenzaussage für diese exakte kleinste Quadrate Interpolation liefert uns der nachstehende Satz.

Satz 3.8 Sei $f \in C([a, b]; \mathbb{R})$ und es sei $(\mathcal{B}_n)_{n \geq n_0}$ eine monotone Folge von Interpolationsknotenverteilungen auf $[a, b]$. Zu $n \geq n_0$ bezeichne $\mathcal{G}_n f$ die exakte kleinste Quadrate Interpolierende zu f , die bezüglich der Interpolationsbasis $\{e_0, e_1\}$ mit $e_k(x) = x^k$ und einem positiven Interpolationsgewicht $w \in C^l(\mathbb{R} \setminus \{0\}; \mathbb{R}_{>0})$ von der Form $w(x) = \|x\|_2^{-\alpha}$ konstruiert wird. Ist nun die Folge $(c(n)r(\mathcal{B}_n)^{\alpha-1})_{n \geq n_0}$ beschränkt, dann konvergiert $\mathcal{G}_n f$ gleichmäßig gegen f auf $[a, b]$ mit $n \rightarrow \infty$, das heißt

$$\lim_{n \rightarrow \infty} \|\mathcal{G}_n f - f\|_\infty = 0. \quad (3.44)$$

Beweis. Für jedes $n \geq n_0$ sei $\mathcal{G}_n = \mathcal{G}_n^+ - \mathcal{G}_n^-$ eine Darstellung des Interpolationsoperators entsprechend Abschnitt 3.7 mit $\mathcal{G}_n^- \rightarrow 0$ für $n \rightarrow \infty$. Zu $\xi \in [a, b]$ sei $\psi_\xi = \psi_\xi(x) = |x - \xi|$. Ferner existiert zu $x \in [a, b]$ ein $M > 0$ mit

$$-2M < f(x) - f(\xi) < 2M$$

aufgrund der Stetigkeit von f und der Kompaktheit von $[a, b]$. Die Stetigkeit von f bei ξ garantiert ferner zu vorgelegtem $\varepsilon > 0$ ein $\delta > 0$ mit

$$|x - \xi| < \delta \implies |f(x) - f(\xi)| < \varepsilon, \quad (3.45)$$

so dass für $x \in [a, b]$ schon

$$-\varepsilon - \frac{2M}{\delta} \psi_\xi(x) \leq f(x) - f(\xi) \leq \varepsilon + \frac{2M}{\delta} \psi_\xi(x) \quad (3.46)$$

unter Beachtung von (3.45) und $\psi_\xi(x)\delta^{-1} \geq 1$ für $|x - \xi| \geq \delta$ folgt. Die Anwendung des positiven und linearen Operators \mathcal{G}_n^+ auf Gleichung (3.46) liefert

$$-\varepsilon \mathcal{G}_n^+ e_0 - \frac{2M}{\delta} \mathcal{G}_n^+ \psi_\xi \leq \mathcal{G}_n^+ f - f(\xi) \mathcal{G}_n^+ e_0 \leq \varepsilon \mathcal{G}_n^+ e_0 + \frac{2M}{\delta} \mathcal{G}_n^+ \psi_\xi$$

beziehungsweise

$$\begin{aligned} -\varepsilon (\mathcal{G}_n + \mathcal{G}_n^-) e_0 - \frac{2M}{\delta} (\mathcal{G}_n + \mathcal{G}_n^-) \psi_\xi &\leq (\mathcal{G}_n + \mathcal{G}_n^-) f - f(\xi) (\mathcal{G}_n + \mathcal{G}_n^-) e_0 \\ &\leq \varepsilon (\mathcal{G}_n + \mathcal{G}_n^-) e_0 + \frac{2M}{\delta} (\mathcal{G}_n + \mathcal{G}_n^-) \psi_\xi. \end{aligned} \quad (3.47)$$

Können wir also

$$\lim_{n \rightarrow \infty} \|\mathcal{G}_n \psi_\xi - \psi_\xi\|_\infty = 0 \quad (3.48)$$

für beliebiges $\xi \in [a, b]$ beweisen, liefert Gleichung (3.47)

$$\lim_{n \rightarrow \infty} |\mathcal{G}_n f(\xi) - f(\xi)| \leq \varepsilon$$

unter Beachtung von $\psi_\xi(\xi) = 0$. Zum Beweis von (3.48) schreiben wir für $x \in [a, b]$

$$\psi_\xi(x) = \psi_\xi^{(1)}(x) + \psi_\xi^{(2)}(x)$$

mit

$$\psi_\xi^{(1)}(x) = \begin{cases} x - \xi & : x \geq \xi \\ 0 & : x < \xi \end{cases}, \quad \psi_\xi^{(2)}(x) = \begin{cases} 0 & : x \geq \xi \\ \xi - x & : x < \xi \end{cases}.$$

Hierfür folgt

$$\begin{aligned}
& \sup_{x \in [a, b]} |\mathcal{G}_n \psi_{\xi}(x) - \psi_{\xi}(x)| \\
&= \sup_{x \in [a, b]} |\mathcal{G}_n(\psi_{\xi}^{(1)}(x) + \psi_{\xi}^{(2)}(x)) - (\psi_{\xi}^{(1)}(x) + \psi_{\xi}^{(2)}(x))| \\
&\leq \sup_{x \in [a, b]} |\mathcal{G}_n \psi_{\xi}^{(2)}(x) - \psi_{\xi}^{(2)}(x)| + \sup_{x \in [a, b]} |\mathcal{G}_n \psi_{\xi}^{(1)}(x) - \psi_{\xi}^{(1)}(x)| \\
&\leq \sup_{x \in [a, \xi]} |\mathcal{G}_n \psi_{\xi}^{(1)}(x) - \psi_{\xi}^{(1)}(x)| + \sup_{x \in [\xi, b]} |\mathcal{G}_n \psi_{\xi}^{(1)}(x) - \psi_{\xi}^{(1)}(x)| \\
&\quad + \sup_{x \in [a, \xi]} |\mathcal{G}_n \psi_{\xi}^{(2)}(x) - \psi_{\xi}^{(2)}(x)| + \sup_{x \in [\xi, b]} |\mathcal{G}_n \psi_{\xi}^{(2)}(x) - \psi_{\xi}^{(2)}(x)| \\
&= \sup_{x \in [a, \xi]} |\mathcal{G}_n \psi_{\xi}^{(1)}(x)| + \sup_{x \in [\xi, b]} |\mathcal{G}_n \psi_{\xi}^{(1)}(x) - \psi_{\xi}^{(1)}(x)| \\
&\quad + \sup_{x \in [a, \xi]} |\mathcal{G}_n \psi_{\xi}^{(2)}(x) - \psi_{\xi}^{(2)}(x)| + \sup_{x \in [\xi, b]} |\mathcal{G}_n \psi_{\xi}^{(2)}(x)|,
\end{aligned}$$

denn $\psi_{\xi}^{(1)}(x) = 0$ für $x \in [a, \xi]$ und $\psi_{\xi}^{(2)}(x) = 0$ für $x \in [\xi, b]$. Ist $B_{n, \xi}^- \cup B_{n, \xi}^+ = \mathcal{B}_n$ eine Zerlegung der unterliegenden Interpolationsknotenverteilung mit $B_{n, \xi}^- \subset [a, \xi)$ und $B_{n, \xi}^+ \subset [\xi, b]$, erhalten wir

$$\begin{aligned}
\sup_{x \in [a, \xi]} |\mathcal{G}_n \psi_{\xi}^{(1)}(x)| &= \sup_{x \in [a, \xi]} \left| \sum_{i=1}^{c(n)} \psi_{\xi}^{(1)}(x_i) \varphi_{i(n)}(x) \right| = \sup_{x \in [a, \xi]} \left| \sum_{i: x_i \in B_{n, \xi}^+} \psi_{\xi}^{(1)}(x_i) \varphi_{i(n)}(x) \right| \\
&= \mathcal{O}(c(n)^2 r(\mathcal{B}_n)^{\alpha+1}) = \mathcal{O}(\eta(n) c(n) r(\mathcal{B}_n)^{\alpha-1}),
\end{aligned}$$

denn für $x \in [a, \xi]$ und Indizes i mit $x_i \in B_{n, \xi}^+$ ist $|\varphi_{i(n)}(x)| = \mathcal{O}(c(n) r(\mathcal{B}_n)^{\alpha+1})$ nach Satz 3.6.

Das Setzen von $\tilde{\psi}_{\xi}^{(1)}(x) = x - \xi$ für $x \in [a, b]$ liefert

$$\begin{aligned}
\tilde{\psi}_{\xi}^{(1)}(x) = \mathcal{G}_n \tilde{\psi}_{\xi}^{(1)}(x) &= \sum_{i: x_i \in B_{n, \xi}^-} \tilde{\psi}_{\xi}^{(1)}(x_i) \varphi_{i(n)}(x) + \sum_{i: x_i \in B_{n, \xi}^+} \tilde{\psi}_{\xi}^{(1)}(x_i) \varphi_{i(n)}(x) \\
&= \sum_{i: x_i \in B_{n, \xi}^-} \tilde{\psi}_{\xi}^{(1)}(x_i) \varphi_{i(n)}(x) + \mathcal{G}_n \psi_{\xi}^{(1)}(x),
\end{aligned}$$

das heißt für beliebiges $x \in [\xi, b]$ gilt

$$\begin{aligned}
|\mathcal{G}_n \psi_{\xi}^{(1)}(x) - \psi_{\xi}^{(1)}(x)| &= \left| \tilde{\psi}_{\xi}^{(1)}(x) - \psi_{\xi}^{(1)}(x) - \sum_{i: x_i \in B_{n, \xi}^-} \tilde{\psi}_{\xi}^{(1)}(x_i) \varphi_{i(n)}(x) \right| \\
&= \left| \sum_{i: x_i \in B_{n, \xi}^-} \tilde{\psi}_{\xi}^{(1)}(x_i) \varphi_{i(n)}(x) \right|.
\end{aligned}$$

Für $x \in [\xi, b]$ und jeden Index i mit $x_i \in B_{n, \xi}^-$ gilt erneut $|\varphi_{i(n)}(x)| = \mathcal{O}(c(n) r(\mathcal{B}_n)^{\alpha+1})$, so dass

$$\sup_{x \in [\xi, b]} |\mathcal{G}_n \psi_{\xi}^{(1)}(x) - \psi_{\xi}^{(1)}(x)| = \mathcal{O}(\eta(n) c(n) r(\mathcal{B}_n)^{\alpha-1})$$

folgt. Mit analogen Argumenten erhalten wir

$$\max \left\{ \sup_{x \in [\xi, b]} \left| \mathcal{G}_n \psi_{\xi}^{(2)}(x) \right|, \sup_{x \in [a, \xi]} \left| \mathcal{G}_n \psi_{\xi}^{(2)}(x) - \psi_{\xi}^{(2)}(x) \right| \right\} = \mathcal{O}(\eta(n)c(n)r(\mathcal{B}_n)^{\alpha-1}),$$

womit insgesamt

$$\sup_{x \in K} \left| \mathcal{G}_n \psi_{\xi}(x) - \psi_{\xi}(x) \right| = \mathcal{O}(\eta(n)c(n)r(\mathcal{B}_n)^{\alpha-1})$$

und $\lim_{n \rightarrow \infty} \|\mathcal{G}_n \psi_{\xi} - \psi_{\xi}\|_{\infty} = 0$ resultiert, was den Beweis beendet. \square

Kapitel 4

Ein Resultat von simultaner Approximation

4.1 Simultane Approximation von konvexen Operatoren	79
4.2 Exkurs in die Theorie der Sobolev-Räume	81
4.3 Simultane Approximation durch kompakte Einbettungen	82

Num die Frage nach simultaner Approximation bei der exakte kleinste Quadrate Interpolation zu motivieren, werden wir zunächst einige klassische Resultate über die *simultane* bzw. *konservative Approximation* bei linearen Operatoren bringen. Im Anschluss geben wir eine kurze Einführung in die Theorie der Sobolev-Räume und beweisen darauf aufbauend ein erstes Resultat von simultaner Approximation für die univariate exakte kleinste Quadrate Interpolation basierend auf dem Theorem von Rellich und Kondrachov über kompakte Einbettungen von Sobolev-Räumen.

4.1 Simultane Approximation von konvexen Operatoren

Zunächst werfen wir einen Blick auf die Definition eines *konvexen Operators* und folgen der Darstellung in [12] basierend auf [25].

Definition 4.1 Sei $f \in C(K; \mathbb{R})$ und $K \subset \mathbb{R}$ ein kompaktes Intervall. Gilt dann für jede m -te dividierte Differenz $[[\xi_0, \dots, \xi_m; f]]$ von f zu $m + 1$ paarweise verschiedenen Punkten $\xi_i \in K$ mit $\xi_0 < \xi_1 < \dots < \xi_m$ die Ungleichung $[[\xi_0, \dots, \xi_m; f]] \geq 0$, so gehört f zur Klasse \mathcal{K}_K^m . Ein Operator $L : C(K; \mathbb{R}) \rightarrow C(K; \mathbb{R})$ heißt *konvex von der Ordnung m* für Werte $m \geq -1$, falls die Implikation

$$f \in \mathcal{K}_K^{m+1} \implies Lf \in \mathcal{K}_K^{m+1}$$

gilt.

Wir stellen fest, dass ein positiver Operator $P : C(K; \mathbb{R}) \rightarrow C(K; \mathbb{R})$ nach obiger Definition konvex von der Ordnung $m = -1$ ist, so dass der Begriff der *Konvexität* von Operatoren als Verallgemeinerung des Begriffs der *Positivität* aufgefasst werden kann. Konvexe Operatoren

werden, vermöge ihrer Definition auf natürliche Weise gegebenen Eigenschaften, als *formerhaltend* bezeichnet.

Sendov und Popov [25] konnten zum Ende der sechziger Jahre des letzten Jahrhunderts, also weit über ein Jahrzehnt nach Veröffentlichung des klassischen Konvergenzsatzes von Bohman und Korovkin, für konvexe Operatoren ein erstes Resultat von simultaner Approximation beweisen.

Satz 4.1 (Sendov, Popov) Seien $L_n : C^k(K; \mathbb{R}) \rightarrow C^k(K; \mathbb{R})$ lineare Operatoren für $n \in \mathbb{N}$, die

$$\lim_{n \rightarrow \infty} \|e_k - L_n e_k\|_\infty = 0, \quad k = 0, 1, 2,$$

für die Monomfunktionen $e_k : x \mapsto x^k$ erfüllen und von der Ordnung $m = -1, 0, 1, \dots, k$ konvex sind. Dann gilt für jedes $f \in C^k(K; \mathbb{R})$ und $0 \leq j \leq k$

$$\lim_{n \rightarrow \infty} \|D^j f - D^j L_n f\|_\infty = 0.$$

Im Laufe der Zeit wurde dieses Ergebnis auf Verallgemeinerungen des Konvexitätsbegriffs erweitert. Hierbei sind beispielsweise Resultate für *fast konvexe Operatoren* von Knoop und Pottinger [12, 13] zu erwähnen. Wir wollen an dieser Stelle nicht allzu detailliert auf diese und weitere Ergebnisse eingehen. Stattdessen bringen wir abschließend ein neueres Resultat von Muñoz-Delgado et al. [20], deren Arbeit im Wesentlichen sämtliche Arbeiten zu dieser Thematik, die auf den formerhaltenden Eigenschaften der Operatoren basieren, verallgemeinern und vereinheitlichen.

Definition 4.2 Die Teilmenge C eines reellen Vektorraums wird *Kegel* genannt, falls für jedes $f \in C$ und $\alpha \in \mathbb{R}_{\geq 0}$ auch $\alpha f \in C$ ist.

Satz 4.2 (Muñoz-Delgado et al.) Sei $K \subset \mathbb{R}^d$ kompakt, $B \subset \mathbb{R}^K$ und $A \subset C(K; \mathbb{R})$ mit $A \subset B$. Ferner bezeichne $L : B \rightarrow \mathbb{R}^K$ einen linearen Operator mit $L(A) \subset C(K; \mathbb{R})$.

Sei weiter $P = \{f \in B \mid Lf \geq 0\}$, C ein Kegel in A und $V \subset A$ ein Unterraum ausgestattet mit den folgenden Eigenschaften:

(1) Es existiert $u \in V$ mit $Lu(x) = 1$ für alle $x \in K$.

(2) Für jedes $z \in K$ existiert ein $\varphi_z \in V \cap C$ mit

$$(2.1) \quad L\varphi_z(z) = 0 < L\varphi_z(x) \text{ für alle } x \in K \setminus \{z\},$$

$$(2.2) \quad \forall f \in A \quad \exists \alpha = \alpha(f) > 0 : \left(\beta \geq \alpha \Rightarrow \beta \varphi_z + f \in C \right).$$

Ist nun $(K_n)_{n \geq 1}$, $K_n : A \rightarrow B$, eine Folge von linearen Operatoren mit $K_n(P \cap C) \subset P$ für $n \geq 1$, und konvergiert $L(K_n f)$ gleichmäßig gegen Lf für jedes $f \in V$, dann konvergiert $L(K_n f)$ gleichmäßig gegen Lf mit $n \rightarrow \infty$ für jedes $f \in A$.

Werfen wir an dieser Stelle noch kurz einen Blick auf Theorem 3.2 für den Spezialfall für Folgen von positiven Operatoren, das heißt den klassischen multivariaten Konvergenzsatz von Bohman und Korovkin. Dieser leitet sich direkt aus Satz 4.2 durch die Wahl von $A = B = C(K; \mathbb{R})$, $C = C(K; \mathbb{R})$, $V = \text{span}\{b^{(1)}, \dots, b^{(k)}\}$ sowie $L = \text{id}_{C(K; \mathbb{R})}$ ab. Die Voraussetzung (1) von Satz 4.2 liefert die Annahme $b^{(1)} \equiv 1$, und unser Begriff der positiven Summierbarkeit spiegelt sich in der Voraussetzung (2.1) von Satz 4.2 wider. Ferner ist Satz 4.2(2.2) aufgrund der simplen Wahl des betrachteten Kegels, das heißt $C = C(K; \mathbb{R})$, trivialerweise erfüllt.

4.2 Exkurs in die Theorie der Sobolev-Räume

Dieser Abschnitt stellt einige grundlegende Resultate über Sobolev-Räume bereit. Das hierbei angestrebte Hauptresultat ist das Theorem von Rellich und Kondrachov über kompakte Einbettungen von Sobolev-Räumen. Aus Gründen der Vollständigkeit haben wir uns dazu entschlossen, diese Resultate explizit in die vorliegende Arbeit aufzunehmen und dabei auf reine Literaturverweise zu verzichten. Wir wollen lediglich die Theorie der L_p -Räume [4] als bekannt voraussetzen. Die gewählte Darstellung orientiert sich an der Monographie [1] vom Adams.

Der Begriff der *schwachen Ableitung* stellt eine natürliche Erweiterung des klassischen Ableitungsbegriffs im Sinne der folgenden Definition dar.

Definition 4.3 Sei $u \in L_{1,loc}(\Omega; \mathbb{R})$ für ein Gebiet $\Omega \subset \mathbb{R}^d$ und $\mu \in \mathbb{N}^d$ ein Multi-Index. Existiert ein $v \in L_{1,loc}(\Omega; \mathbb{R})$ derart, dass

$$\int_{\Omega} u D^{\mu} \psi \, dx = (-1)^{|\mu|} \int_{\Omega} v \psi \, dx$$

für alle Testfunktionen $\psi \in C_c^{\infty}(\Omega; \mathbb{R})$ gilt, so heißt v die *schwache μ -Ableitung* von u , die als $D^{\mu}u = v$ notiert wird.

Schwache Ableitungen sind, sofern diese existieren, eindeutig bestimmt und stimmen mit den entsprechenden klassischen Ableitungen überein, falls die betrachtete Funktion insbesondere im klassischen Sinn differenzierbar ist.

Definition 4.4 Sei $\Omega \subset \mathbb{R}^d$ ein Gebiet und $1 \leq k < \infty$ sowie $1 \leq p \leq \infty$.

- (a) Mit $W_p^k(\Omega)$ wird der Raum aller $u \in L_p(\Omega)$ bezeichnet, die $D^{\mu}u \in L_p(\Omega)$ für alle Multi-Indizes $|\mu| \leq k$ erfüllen.
- (b) Mit $\hat{W}_p^k(\Omega)$ wird die Vervollständigung des Raums $C_c^{\infty}(\Omega)$ bezüglich der *Sobolev-Norm*

$$\|u\|_{W_p^k(\Omega)} = \left(\sum_{0 \leq |\mu| \leq k} \|D^{\mu}u\|_{L_p(\Omega)}^p \right)^{\frac{1}{p}}, \quad u \in W_p^k(\Omega), \quad (4.1)$$

bezeichnet.

Die Räume $W_p^k(\Omega)$ und $\hat{W}_p^k(\Omega)$ heißen *Sobolev-Räume*.

Satz 4.3 Sei $\Omega \subset \mathbb{R}^d$ ein Gebiet und $W_p^k(\Omega)$ ein Sobolev-Raum. Dann ist dieser vollständig bezüglich der Sobolev-Norm (4.1), das heißt $W_p^k(\Omega)$ ist ein Banach-Raum. Die Räume $W_2^k(\Omega)$ und $\hat{W}_2^k(\Omega)$ sind Hilberträume bezüglich des inneren Produktes

$$(u, v)_{W_2^k(\Omega)} = \sum_{0 \leq |\mu| \leq k} (D^{\mu}u, D^{\mu}v)_{L_2(\Omega)}, \quad u, v \in W_2^k(\Omega).$$

Satz 4.4 (Meyers, Serrin) Es gilt die Identität

$$W_p^k(\Omega) = H_p^k(\Omega),$$

wobei mit $H_p^k(\Omega)$ die Vervollständigung der Menge

$$\left\{ u \in C^{\infty}(\Omega) \mid \|u\|_{W_p^k(\Omega)} < \infty \right\}$$

bezüglich der Sobolev-Norm $\|\cdot\|_{W_p^k(\Omega)}$ bezeichnet wird.

Definition 4.5 Die Menge $\Omega \subset \mathbb{R}^d$ genügt der *Kegel-Bedingung*, falls es einen endlichen Kegel $C \subsetneq \mathbb{R}^d$ gibt, so dass jedes $x \in \Omega$ die Spitze eines endlichen Kegels $C_x \subset \Omega$ ist, der kongruent (das heißt deckungsgleich) zu C ist. Hierbei wird ein Kegel endlich genannt, falls C dieser der Schnitt eines Kegels nach Definition 4.2 mit einer Ursprungskugel mit endlichem Radius ist.

Definition 4.6 Seien $(X, \|\cdot\|_X)$ und $(Y, \|\cdot\|_Y)$ Banach-Räume. Dann wird der Raum X *kompakt eingebettet* in Y genannt, notiert als $X \hookrightarrow Y$, falls $X \subset Y$ und eine Konstante $C > 0$ derart existiert, so dass

$$\|x\|_Y \leq C \|x\|_X, \quad x \in X,$$

und jede beschränkte Folge in X folgenkompakt in Y ist.

Theorem 4.1 (Rellich, Kondrachov) Sei $\Omega \subset \mathbb{R}^d$ ein Gebiet und Ω_b ein beschränktes Teilgebiet von Ω . Seien ferner k, m und p natürliche Zahlen mit $k \geq 1$ und $1 \leq p < \infty$. Erfüllt Ω die Kegel-Bedingung nach Definition 4.5 und ist $mp > d$, dann sind die Einbettungen

$$W_p^{k+m}(\Omega) \hookrightarrow C_b^m(\Omega_b)$$

und

$$\dot{W}_p^{k+m}(\Omega) \hookrightarrow C_b^m(\Omega_b)$$

kompakt.

Hierbei wird mit $C_b^m(\Omega)$ der Raum aller $u \in C^m(\Omega)$ bezeichnet, so dass $D^\mu u$ beschränkt auf Ω für alle $0 \leq |\mu| \leq m$ ist. Mit der Norm

$$\|u\|_{C_b^m(\Omega)} = \max_{0 \leq |\mu| \leq m} \|D^\mu u\|_\infty$$

wird $C_b^m(\Omega)$ zum Banach-Raum.

Es sei angemerkt, dass Theorem 4.1 auch für andere Werte von m, p und d mit entsprechenden Einbettungen Gültigkeit besitzt. Hierfür verweisen wir auf [1]. Unter Zuhilfenahme der hier angegebenen Version von Theorem 4.1 ist es uns möglich, eine erste Aussage über die simultane Approximation von exakte kleinste Quadrate Interpolierenden zu treffen.

4.3 Simultane Approximation durch kompakte Einbettungen

Den Resultaten dieses Abschnitts liegen die folgenden Annahmen zu Grunde: Wir betrachten Funktionen

$$f \in C^2([a, b]; \mathbb{R}) \cap \dot{W}_2^2(a, b)$$

für ein Intervall $K = [a, b] \subset \mathbb{R}$. Ist dann $w \in C^l(\mathbb{R} \setminus \{0\}; \mathbb{R}_{>0})$ ein positives Interpolationsgewicht und sind $b^{(1)}, \dots, b^{(k)} \in C^m(K; \mathbb{R})$ Interpolationsbasisfunktionen mit $l, m \geq 2$, gilt für die exakte kleinste Quadrate Interpolierende zu f bezüglich w , der $b^{(i)}$ und einer Interpolationsknotenverteilung \mathcal{B}_n schon $\mathcal{G}_n f \in C^2(K; \mathbb{R})$ nach Theorem 1.3.

Wir bringen nun ein erstes Resultat über simultane Approximation.

Satz 4.5 Sei $f \in C^2([a, b]; \mathbb{R}) \cap \dot{W}_2^2(a, b)$, $w \in C^l(\mathbb{R} \setminus \{0\}; \mathbb{R}_{>0})$ ein positives Interpolationsgewicht von der Form $w(x) = \|x\|_2^{-\alpha}$ und die Menge $\{b^{(1)}, \dots, b^{(k)}\} \subset C^m(K; \mathbb{R})$ eine konstante, lineare oder quadratische monomiale Interpolationsbasis mit $l, m \geq 2$. Für festes $n_0 \in \mathbb{N}$ bezeichne $(\mathcal{B}_n)_{n \geq n_0}$ eine monotone Folge von Interpolationsknotenverteilungen in K . Es gelte ferner

$$\lim_{n \rightarrow \infty} \|\mathcal{G}_n e_k - e_k\|_\infty = 0$$

für die drei Monomfunktionen $e_k : x \mapsto x^k$, $0 \leq k \leq 2$. Dann gelten für genügend großes α die folgenden Aussagen:

- (1) $\mathcal{G}_n f$ konvergiert schwach gegen f in $W_2^1(a, b)$ mit $n \rightarrow \infty$.
- (2) $\mathcal{G}_n f$ konvergiert schwach gegen f in $\dot{W}_2^2(a, b)$ mit $n \rightarrow \infty$.

Beweis. Zunächst einmal zum Beweis von (1). Nach Satz 4.4 genügt es zu zeigen, dass

$$(\mathcal{G}_n f, v)_{W_2^1(a, b)} \longrightarrow (f, v)_{W_2^1(a, b)}$$

mit $n \rightarrow \infty$ für beliebiges $v \in C^\infty(a, b)$ gilt. Dies folgt jedoch direkt aus

$$(\mathcal{G}_n f, v)_{L_2(a, b)} \longrightarrow (f, v)_{L_2(a, b)}, \quad n \rightarrow \infty,$$

nach Satz 3.2, Theorem 3.3, Satz 3.8 sowie

$$\begin{aligned} (D\mathcal{G}_n f, Dv)_{L_2(a, b)} &= \mathcal{G}_n f Dv \Big|_a^b - (\mathcal{G}_n f, D^2 v)_{L_2(a, b)} \\ &\longrightarrow f Dv \Big|_a^b - (f, D^2 v)_{L_2(a, b)} = (Df, Dv)_{L_2(a, b)}, \quad n \rightarrow \infty, \end{aligned}$$

durch partielle Integration, wobei wir der üblichen Notation folgend $g \Big|_a^b := g(b) - g(a)$ gesetzt haben. Für den Beweis von (2) reicht es unter Beachtung von (1) und der Definition 4.4

$$(D^2 \mathcal{G}_n f, D^2 v)_{L_2(a, b)} \longrightarrow (D^2 f, D^2 v)_{L_2(a, b)}, \quad n \rightarrow \infty,$$

für jedes $v \in C_c^\infty(a, b)$ zu zeigen. Doch dieses liefert unmittelbar die Rechnung

$$\begin{aligned} (D^2 \mathcal{G}_n f, D^2 v)_{L_2(a, b)} &= D\mathcal{G}_n f D^2 v \Big|_a^b - (D\mathcal{G}_n f, D^3 v)_{L_2(a, b)} \\ &= D\mathcal{G}_n f D^2 v \Big|_a^b - \mathcal{G}_n f D^3 v \Big|_a^b + (\mathcal{G}_n f, D^4 v)_{L_2(a, b)} = (\mathcal{G}_n f, D^4 v)_{L_2(a, b)} \\ &\longrightarrow (f, D^4 v)_{L_2(a, b)} = \dots = (D^2 f, D^2 v)_{L_2(a, b)}, \quad n \rightarrow \infty. \end{aligned}$$

□

Aus Satz 4.5(1) lässt sich direkt das nachstehende Korollar ableiten.

Korollar 4.1 Sei $f \in C^2([a, b]; \mathbb{R}) \cap \dot{W}_2^2(a, b)$ und $\mathcal{G}_n f \in C^2(K; \mathbb{R})$ für $n \geq n_0$ wie in Satz 4.5. Dann existiert zu jedem $\varepsilon > 0$ ein Index $n_1 \in \mathbb{N}$ und $\alpha_{n_0}, \dots, \alpha_{n_1} \in \mathbb{R}_{\geq 0}$ mit $\sum_{j=n_0}^{n_1} \alpha_j = 1$ derart, dass

$$\left\| Df - \sum_{j=n_0}^{n_1} \alpha_j D\mathcal{G}_j f \right\|_{L_2(a, b)} < \varepsilon$$

gilt.

Beweis. Die Behauptung folgt direkt aus dem Satz von Mazur [29] über schwach konvergente Folgen in Banach-Räumen. \square

Als Konsequenz von Satz 4.5(2) erhalten wir das folgende Resultat:

Theorem 4.2 *Zu $f \in C^2([a, b]; \mathbb{R}) \cap \mathring{W}_2^2(a, b)$ bezeichne $\mathcal{G}_n f \in C^2(K; \mathbb{R})$ für $n \geq n_0$ die dazugehörigen exakte kleinste Quadrate Interpolierende wie in Satz 4.5. Dann existiert eine Teilfolge $(n_j)_{j \in \mathbb{N}} \subset \mathbb{N}$ von Indizes derart, dass $\mathcal{G}_{n_j} f$ gegen f in $C_b^1((a, b); \mathbb{R})$ mit $j \rightarrow \infty$ konvergiert, das heißt es gilt*

$$\lim_{j \rightarrow \infty} \|\mathcal{G}_{n_j} f - f\|_{C_b^1((a, b); \mathbb{R})} = 0.$$

Beweis. Offensichtlich erfüllt $\text{int} K = (a, b)$ die Kegel-Bedingung nach Definition 4.5, so dass wir Theorem 4.1 anwenden können. Mit $\Omega = \Omega_b = \text{int} K$, $d = 1$, $p = 2$, $m = 1$ und $k = 1$ folgt

$$\mathring{W}_2^2(a, b) \hookrightarrow C_b^1(a, b),$$

also die kompakte Einbettung des Sobolev-Raumes $\mathring{W}_2^2(a, b)$ in den Raum $C_b^1(a, b)$. Aufgrund der schwachen Konvergenz von $\mathcal{G}_n f$ gegen f , ist die Folge $(\mathcal{G}_n f)_{n \geq n_0}$ beschränkt in $\mathring{W}_2^2(a, b)$, so dass wir infolge der kompakten Einbettung eine Teilfolge $(n_j)_{j \in \mathbb{N}}$ mit

$$\lim_{j \rightarrow \infty} \|\mathcal{G}_{n_j} f - f\|_{C_b^1(a, b)} = 0$$

auswählen können. \square

Korollar 4.2 *Zu $f \in C^2([a, b]; \mathbb{R}) \cap \mathring{W}_2^2(a, b)$ bezeichne $\mathcal{G}_n f \in C^2(K; \mathbb{R})$ für $n \geq n_0$ die dazugehörigen exakte kleinste Quadrate Interpolierende wie in Satz 4.5. Dann existiert eine Teilfolge $(n_j)_{j \in \mathbb{N}} \subset \mathbb{N}$ von Indizes mit*


$$\lim_{j \rightarrow \infty} \sup_{x \in (a, b)} |D\mathcal{G}_{n_j} f(x) - Df(x)| = 0.$$

Beweis. Theorem 4.2. \square

Kapitel 5

Anwendungen im Kontext der numerischen Approximation von Erhaltungsgleichungen

5.1	Konservative Verfahren für die nichtlineare Advektionsgleichung	85
5.2	Ein konservatives Verfahren aus exakte kleinste Quadrate Interpolierenden	87
5.3	Ein informeller Exkurs über Upwind-Verfahren	94
5.4	Ad hoc Upwind-Verfahren durch Variation der Parameter des Interpolationsgewichtes	98
5.5	Numerische Beispiele	99

 heorem 2.1 hat die Grundlagen zur Konstruktion von Finite Differenzen-Verfahren zur numerischen Approximation von nichtlinearen Advektionsgleichungen gelegt, welche wir innerhalb dieses Kapitels speziell auf Basis der exakte kleinste Quadrate Interpolation konstruieren werden. Zunächst führen wir den Begriff der *schwachen Lösung* einer nichtlinearen Advektionsgleichung ein und zeigen dann, wie durch Ableitungs-Approximationen ähnlich denen in Abschnitt 2.4 konservative Finite Differenzen-Verfahren gewonnen werden können. Nach einem Exkurs über die Notwendigkeit von Upwind-Verfahren, bringen wir diese Arbeit abschließend numerische Beispiele.

5.1 Konservative Verfahren für die nichtlineare Advektionsgleichung

Eine zentrale Rolle bei der numerischen Approximation des Cauchy-Problems für die nichtlineare Advektionsgleichung

$$u(x, t)_t + f(u(x, t))_x = 0, \quad x \in \mathbb{R}, \quad t > 0, \quad (5.1a)$$

$$u(x, 0) = u_0(x), \quad x \in \mathbb{R}, \quad (5.1b)$$

spielen die so genannten *konservativen Verfahren*. Ziel dieses Abschnitts ist die Konstruktion konservativer Verfahren mit Hilfe von Ableitungs-Approximationen aus exakte kleinste Quadrate

Interpolierenden. Wir nehmen von nun an für (5.1a) eine differenzierbare Flussfunktion f an. Wir interessieren uns an dieser Stelle nicht für klassische, das heißt differenzierbare Lösungen des Cauchy-Problems (5.1a,b), sondern für die Approximation von schwachen Lösungen bzw. den gasdynamisch relevanten Entropie-Lösungen.

Definition 5.1 (a) Eine Funktion $u \in L_{1,\text{loc}}(\mathbb{R} \times \mathbb{R}_{>0})$ heißt *schwache Lösung* des Cauchy-Problems (5.1a,b) für die nichtlineare Advektionsgleichung, falls

$$\int_0^\infty \int_{-\infty}^\infty \varphi_t(x,t)u(x,t) + \varphi_x(x,t)f(u(x,t)) dx dt = - \int_{-\infty}^\infty \varphi(x,0)u(x,0) dx$$

für alle Testfunktionen $\varphi \in C_c^\infty(\mathbb{R} \times \mathbb{R}_{>0})$ gilt.

(b) Zu gegebener Flussfunktion f in (5.1a) seien ψ und η differenzierbare Funktionen mit

$$\psi'(u) = \eta'(u)f'(u).$$

Dann heißt ψ *Entropie-Fluss* und η *Entropie-Funktion*. Ein $u \in L_{1,\text{loc}}(\mathbb{R} \times \mathbb{R}_{>0})$ heißt *Entropie-Lösung* von (5.1a,b), falls die Ungleichung

$$\eta(u)_t + \psi(u)_x \leq 0 \quad (5.2)$$

im schwachen Sinn gilt, das heißt

$$\int_0^\infty \int_{-\infty}^\infty \varphi_t(x,t)\eta(u(x,t)) + \varphi_x(x,t)\psi(u(x,t)) dx dt \leq - \int_{-\infty}^\infty \varphi(x,0)\eta(u(x,0)) dx$$

für alle *positiven* Testfunktionen $\varphi \in C_c^\infty(\mathbb{R} \times \mathbb{R}_{>0})$.

Aus der Theorie der numerischen Approximation von hyperbolischen Differentialgleichungen [18, 9, 7] ist bekannt, dass konservative Verfahren, die gleichzeitig konsistent zur unterliegenden Differentialgleichung sind, im Fall der L_1 -Konvergenz nach dem Satz von Lax und Wendroff schon gegen eine schwache Lösung bzw. eine Entropie-Lösung konvergieren. Letzteres trifft zu, falls eine diskrete Entropie-Ungleichung in Analogie zu (5.2) erfüllt ist. Wir kommen nun zur Definition der Begriffe *Konsistenz* und *Konservativität* von Verfahren.

Wenden wir uns hierfür zunächst der Diskretisierung des Cauchy-Problems (5.1a,b) zu. Aus Gründen der Einfachheit nutzen wir eine Euler-Diskretisierung der Zeitableitung in (5.1a), die bekannterweise konsistent von erster Ordnung ist. Hierfür sei $\Delta t > 0$ fest gewählt und es bezeichne $U^n(x)$ den Wert

$$U^n(x) = u(x, n\Delta t), \quad n \in \mathbb{N},$$

für eine (klassische) Lösung u der nichtlinearen Advektionsgleichung. Somit geht (5.1a,b) in die semi-diskretisierte Form

$$U^{n+1}(x) = U^n(x) - \Delta t f(u(x, n\Delta t))_x \quad (5.3)$$

über. Für die Behandlung der räumlichen Ableitung $f(u(x, n\Delta t))_x$ ist es im Weiteren mit Blick auf die obige Konvergenzaussage wünschenswert, dass die daraus resultierende vollständige Diskretisierung ein konservatives Verfahren im Sinne der folgenden Definition liefert.

Definition 5.2 Für feste $p, q \in \mathbb{N}$ wird eine Diskretisierung der nichtlinearen Advektionsgleichung (5.1a) von der Form

$$U_j^{n+1} = U_j^n - \frac{\Delta t}{\Delta x} \left(F(U_{j-p}^n, U_{j-p+1}^n, \dots, U_{j+q}^n) - F(U_{j-p-1}^n, U_{j-p}^n, \dots, U_{j+q-1}^n) \right) \quad (5.4)$$

als *konservativ* bezeichnet. Bei dem hieraus resultierenden numerischen Verfahren spricht man von einem *konservativen Verfahren*. F heißt *numerische Flussfunktion*.

Für konservative Verfahren ist im Kontext des Konvergenzsatzes von Lax und Wendroff die Begriffsbildung der *Konsistenz* zur unterliegenden Differentialgleichung essentiell.

Definition 5.3 Das konservative Verfahren (5.4) heißt *konsistent* zur nichtlinearen Advektionsgleichung (5.1a), falls

$$F(\tilde{u}, \dots, \tilde{u}) = f(\tilde{u})$$

für alle $\tilde{u} \in \mathbb{R}$ gilt. Existiert weiter für $\tilde{u} \in \mathbb{R}$ eine Konstante $C_{\tilde{u}} \geq 0$ mit

$$|F(v_1, \dots, v_{p+q+1}) - f(\tilde{u})| \leq C_{\tilde{u}} \max_{1 \leq i \leq p+q+1} |v_i - \tilde{u}|, \quad (5.5)$$

so wird F Lipschitz-stetig bei \tilde{u} genannt.

Die punktweise Lipschitz-Stetigkeit (5.5) der numerischen Flussfunktion impliziert unmittelbar deren Konsistenz zur Advektionsgleichung, das heißt $F(\tilde{u}, \dots, \tilde{u}) = f(\tilde{u})$.

Zur Vereinfachung der Notation schreiben wir für (5.4) im Folgenden auch $U_j^{n+1} = \mathcal{H}_{\Delta t}(U^n; j)$ mit

$$\mathcal{H}_{\Delta t}(U^n; j) = U_j^n - \frac{\Delta t}{\Delta x} \left(F(U^n; j) - F(U^n; j-1) \right)$$

sowie $F(U^n; j) = F(U_{j-p}^n, U_{j-p+1}^n, \dots, U_{j+q}^n)$.

5.2 Ein konservatives Verfahren aus exakte kleinste Quadrate Interpolierenden

In Abschnitt 2.4 haben wir gezeigt, dass durch die Differentiation der exakte kleinste Quadrate Interpolierenden $\mathcal{G}f$ zu einer Funktion f , deren Konstruktion auf einer linearen monomialen Interpolationsbasis sowie einem passend Träger-eingeschränkten unsymmetrischen und ungewichteten Norm-inversen positiven Interpolationsgewicht

$$w \equiv w[\alpha|\beta](x) = \begin{cases} |x|^{-\alpha} & : x < 0 \\ |x|^{-\beta} & : x > 0 \end{cases}, \quad \alpha, \beta \in \mathbb{N}$$

basiert, eine Finite Differenz der Form

$$\frac{d}{dx} \mathcal{G}f(x_j) \sim \left[-\frac{1}{2\Delta x} \left(1 + \frac{\Delta x^\alpha - \Delta x^\beta}{\Delta x^\alpha + \Delta x^\beta} \right) \right] \bullet \bullet \left[\frac{1}{\Delta x} \frac{\Delta x^\alpha - \Delta x^\beta}{\Delta x^\alpha + \Delta x^\beta} \right] \bullet \bullet \left[\frac{1}{2\Delta x} \left(1 - \frac{\Delta x^\alpha - \Delta x^\beta}{\Delta x^\alpha + \Delta x^\beta} \right) \right], \quad (5.6)$$

gewonnen werden kann. Verwenden wir diese Finite Differenz zur Approximation der räumlichen Ableitung in der semi-diskretisierten Form der nichtlinearen Advektionsgleichung (5.3), resultiert die vollständig diskretisierte Version hiervon durch

$$\begin{aligned} U_j^{n+1} = U_j^n - \frac{\Delta t}{\Delta x} & \left(-\frac{1}{2} \left(1 + \frac{\Delta x^\alpha - \Delta x^\beta}{\Delta x^\alpha + \Delta x^\beta} \right) f(U_{j-1}^n) + \dots \right. \\ & \left. \dots + \frac{\Delta x^\alpha - \Delta x^\beta}{\Delta x^\alpha + \Delta x^\beta} f(U_j^n) + \frac{1}{2} \left(1 - \frac{\Delta x^\alpha - \Delta x^\beta}{\Delta x^\alpha + \Delta x^\beta} \right) f(U_{j+1}^n) \right). \end{aligned} \quad (5.7)$$

Wie uns nachstehender Satz lehrt, ist das Verfahren (5.7) sowohl konservativ als auch konsistent zur nichtlinearen Advektionsgleichung im Sinne der Definition 5.3.

Satz 5.1 *Durch (5.7) ist ein konservatives Verfahren zur numerischen Behandlung der nichtlinearen Advektionsgleichung (5.1a) gegeben. Das Verfahren ist ferner konsistent im Sinne der Definition 5.3.*

Beweis. Mit Hilfe der numerischen Flussfunktion $F = F(u, v)$, definiert durch

$$F(u, v) = \frac{1}{2} \left(1 + \frac{\Delta x^\alpha - \Delta x^\beta}{\Delta x^\alpha + \Delta x^\beta} \right) f(u) + \frac{1}{2} \left(1 - \frac{\Delta x^\alpha - \Delta x^\beta}{\Delta x^\alpha + \Delta x^\beta} \right) f(v),$$

kann (5.7) in der Form

$$U_j^{n+1} = U_j^n - \frac{\Delta t}{\Delta x} \left(F(U_j^n, U_{j+1}^n) - F(U_{j-1}^n, U_j^n) \right)$$

geschrieben werden. Demnach ist das Verfahren konservativ. Die Konsistenz folgt direkt unter Berücksichtigung von $f \in C^1(\mathbb{R}; \mathbb{R})$ durch die Abschätzungen

$$\begin{aligned} & |F(v_1, v_2) - f(\tilde{u})| \\ &= \left| \frac{1}{2} \left(1 + \frac{\Delta x^\alpha - \Delta x^\beta}{\Delta x^\alpha + \Delta x^\beta} \right) f(v_1) - \frac{1}{2} f(\tilde{u}) + \frac{1}{2} \left(1 + \frac{\Delta x^\alpha - \Delta x^\beta}{\Delta x^\alpha + \Delta x^\beta} \right) f(v_2) - \frac{1}{2} f(\tilde{u}) \right| \\ &= \left| \frac{1}{2} (f(v_1) - f(\tilde{u})) + \frac{1}{2} (f(v_2) - f(\tilde{u})) + \frac{1}{2} \frac{\Delta x^\alpha - \Delta x^\beta}{\Delta x^\alpha + \Delta x^\beta} (f(v_1) - f(v_2)) \right| \\ &\leq \frac{1}{2} \|f'\|_\infty |v_1 - \tilde{u}| + \frac{1}{2} \|f'\|_\infty |v_2 - \tilde{u}| + \frac{1}{2} \left| \frac{\Delta x^\alpha - \Delta x^\beta}{\Delta x^\alpha + \Delta x^\beta} \right| \|f'\|_\infty |v_1 - v_2| \\ &\leq \frac{1}{2} \|f'\|_\infty (|v_1 - \tilde{u}| + |v_2 - \tilde{u}|) + \frac{1}{2} \left| \frac{\Delta x^\alpha - \Delta x^\beta}{\Delta x^\alpha + \Delta x^\beta} \right| \|f'\|_\infty (|v_1 - \tilde{u}| + |v_2 - \tilde{u}|) \\ &\leq \left(1 + \left| \frac{\Delta x^\alpha - \Delta x^\beta}{\Delta x^\alpha + \Delta x^\beta} \right| \right) \|f'\|_\infty \max_{1 \leq i \leq 2} |v_i - \tilde{u}|. \end{aligned} \quad (5.8)$$

□

Bevor das Verfahren (5.7) an einigen Fallbeispielen näher untersucht wird, konstruieren wir zunächst 5-Punkte Finite Differenzen zur Approximation des Wertes $\frac{d}{dx}\mathcal{G}f(x_j)$ in Analogie zu Abschnitt 2.4.

Wir beginnen mit der Konstruktion einer 5-Punkte Finite Differenz auf Basis einer exakten kleinsten Quadrate Interpolierenden mit zugrundeliegender linearer monomialer Interpolationsbasis.

Zu $x_j \in \{x_3, \dots, x_{n-2}\}$ ist dann $\underline{b}(x_{j+i}) = (1, x_j + i\Delta x)^T$ für Indizes $i \in \{0, \pm 1, \pm 2\}$ und $\frac{d}{dx}\underline{b}(x_{j+i}) = (0, 1)^T$. Mit

$$\begin{aligned} w^{(j-2)}(x_j) &= |x_j - x_{j-2}|^{-\beta} = (2\Delta x)^{-\beta}, & w^{(j-1)}(x_j) &= |x_j - x_{j-1}|^{-\beta} = \Delta x^{-\beta}, \\ w^{(j+1)}(x_j) &= |x_j - x_{j+1}|^{-\beta} = \Delta x^{-\alpha}, & w^{(j+2)}(x_j) &= |x_j - x_{j+2}|^{-\beta} = (2\Delta x)^{-\alpha}, \end{aligned}$$

erhalten wir die Matrix $R(x_j) = \begin{pmatrix} R_j^{[11]} & R_j^{[12]} \\ R_j^{[21]} & R_j^{[22]} \end{pmatrix}$ mit den Einträgen

$$\begin{aligned} R_j^{[11]} &= w^{(j-2)}(x_j) + w^{(j-1)}(x_j) + w^{(j+1)}(x_j) + w^{(j+2)}(x_j), \\ R_j^{[12]} = R_j^{[21]} &= x_{j-2}w^{(j-2)}(x_j) + x_{j-1}w^{(j-1)}(x_j) + x_{j+1}w^{(j+1)}(x_j) + x_{j+2}w^{(j+2)}(x_j), \\ R_j^{[22]} &= x_{j-2}^2w^{(j-2)}(x_j) + x_{j-1}^2w^{(j-1)}(x_j) + x_{j+1}^2w^{(j+1)}(x_j) + x_{j+2}^2w^{(j+2)}(x_j). \end{aligned}$$

Die zugehörige Finite Differenz ist nach Theorem 2.1 durch

$$\frac{d}{dx}\mathcal{G}f(x_j) \sim \boxed{\varphi'_{j-2}(x_j)} \bullet \bullet \boxed{\varphi'_{j-1}(x_j)} \bullet \bullet \boxed{\varphi'_j(x_j)} \bullet \bullet \boxed{\varphi'_{j+1}(x_j)} \bullet \bullet \boxed{\varphi'_{j+2}(x_j)} \quad (5.9)$$

mit den Werten

$$\begin{aligned} \varphi'_{j-2}(x_j) &= -\frac{1}{\Delta x} \frac{2^{1+\alpha}\Delta x^\alpha}{2^\alpha(4+2^\beta)\Delta x^\alpha + 2^\beta(4+2^\alpha)\Delta x^\beta}, \\ \varphi'_{j-1}(x_j) &= -\frac{1}{\Delta x} \frac{2^{\alpha+\beta}\Delta x^\alpha}{2^\alpha(4+2^\beta)\Delta x^\alpha + 2^\beta(4+2^\alpha)\Delta x^\beta}, \\ \varphi'_j(x_j) &= \frac{1}{\Delta x} \frac{2^\alpha(2+2^\beta)\Delta x^\alpha - 2^\beta(2+2^\alpha)\Delta x^\beta}{2^\alpha(4+2^\beta)\Delta x^\alpha + 2^\beta(4+2^\alpha)\Delta x^\beta}, \\ \varphi'_{j+1}(x_j) &= \frac{1}{\Delta x} \frac{2^{\alpha+\beta}\Delta x^\beta}{2^\alpha(4+2^\beta)\Delta x^\alpha + 2^\beta(4+2^\alpha)\Delta x^\beta}, \\ \varphi'_{j+2}(x_j) &= \frac{1}{\Delta x} \frac{2^{1+\beta}\Delta x^\beta}{2^\alpha(4+2^\beta)\Delta x^\alpha + 2^\beta(4+2^\alpha)\Delta x^\beta}, \end{aligned}$$

gegeben.

Eine Konsistenzanalyse von (5.9) analog zu der des Ausdrucks (5.6) in Kapitel 2 basiert auf einer Taylor-Entwicklung der Form

$$\underbrace{\frac{d}{dx} \mathcal{G} f(x_j)}_{\approx \frac{d}{dx} f(x_j)} - \left(\varphi'_{j-2}(x_j) f_{j-2} + \varphi'_{j-1}(x_j) f_{j-1} + \varphi'_j(x_j) f_j + \varphi'_{j+1}(x_j) f_{j+1} + \varphi'_{j+2}(x_j) f_{j+2} \right) =$$

$$+ \frac{\Delta x}{2!} \frac{\mathfrak{D}_{\alpha,\beta}^3}{\mathfrak{D}_{\alpha,\beta}^2} \frac{d^2}{dx^2} f_j - \frac{\Delta x^2}{3!} \frac{\mathfrak{D}_{\alpha,\beta}^4}{\mathfrak{D}_{\alpha,\beta}^2} \frac{d^3}{dx^3} f_j + \frac{\Delta x^3}{4!} \frac{\mathfrak{D}_{\alpha,\beta}^5}{\mathfrak{D}_{\alpha,\beta}^2} \frac{d^4}{dx^4} f_j$$

$$- \frac{\Delta x^4}{5!} \frac{\mathfrak{D}_{\alpha,\beta}^6}{\mathfrak{D}_{\alpha,\beta}^2} \frac{d^5}{dx^5} f_j + \frac{\Delta x^5}{6!} \frac{\mathfrak{D}_{\alpha,\beta}^7}{\mathfrak{D}_{\alpha,\beta}^2} \frac{d^6}{dx^6} f_j + \mathcal{O}(\Delta x^6), \quad (5.10)$$

wobei wir

$$\mathfrak{D}_{\alpha,\beta}^n := 2^\alpha (2^n + 2^\beta) \Delta x^\alpha + (-1)^n 2^\beta (2^n + 2^\alpha) \Delta x^\beta, \quad n \in \mathbb{N}, \quad (5.11)$$

gesetzt haben. Hierbei hängt die Konsistenzordnung von (5.9) mit Blick auf (5.10) von dem Verhalten der Terme $\mathfrak{D}_{\alpha,\beta}^n / \mathfrak{D}_{\alpha,\beta}^2$ ab. Im Fall $n = 2$ ist sofort einzusehen, dass

$$\mathfrak{D}_{\alpha,\beta}^2 = 2^\alpha (2^2 + 2^\beta) \Delta x^\alpha + 2^\beta (2^2 + 2^\alpha) \Delta x^\beta > 0$$

für alle $\alpha, \beta \in \mathbb{N}$ gilt. Daher ist $\mathfrak{D}_{\alpha,\beta}^n / \mathfrak{D}_{\alpha,\beta}^2 \in \mathbb{R}$ für alle $n \in \mathbb{N}$. Für $\alpha = \beta$ erhalten wir

$$\frac{\mathfrak{D}_{\alpha,\beta}^n}{\mathfrak{D}_{\alpha,\beta}^2} = \begin{cases} 0 & : n \text{ ungerade} \\ \frac{2^n + 2^\alpha}{2^2 + 2^\alpha} & : n \text{ gerade} \end{cases}$$

und $\alpha \neq \beta$ impliziert

$$\frac{\mathfrak{D}_{\alpha,\beta}^n}{\mathfrak{D}_{\alpha,\beta}^2} = 0 \iff 2^\alpha (2^n + 2^\beta) \Delta x^\alpha + (-1)^n 2^\beta (2^n + 2^\alpha) \Delta x^\beta = 0.$$

Das heißt, (5.9) ist konsistent von erster Ordnung falls $2^\alpha (2^3 + 2^\beta) \Delta x^\alpha - 2^\beta (2^3 + 2^\alpha) \Delta x^\beta \neq 0$ und konsistent von zweiter Ordnung, falls $\alpha = \beta$ oder $2^\alpha (2^3 + 2^\beta) \Delta x^\alpha - 2^\beta (2^3 + 2^\alpha) \Delta x^\beta = 0$ gilt.

Wir kommen nun zur Konstruktion einer 5-Punkte Finite Differenz, wobei für die exakte kleinste Quadrate Interpolierende eine lineare monomiale Interpolationsbasis angenommen wird.

Hierfür betrachten wir die Matrix

$$R(x_j) = \begin{pmatrix} R_j^{[11]} & R_j^{[12]} & R_j^{[13]} \\ R_j^{[21]} & R_j^{[22]} & R_j^{[23]} \\ R_j^{[31]} & R_j^{[32]} & R_j^{[33]} \end{pmatrix}$$

mit den Einträgen

$$\begin{aligned}
R_j^{[11]} &= w^{(j-2)}(x_j) + w^{(j-1)}(x_j) + w^{(j+1)}(x_j) + w^{(j+2)}(x_j), \\
R_j^{[12]} = R_j^{[21]} &= x_{j-2}w^{(j-2)}(x_j) + x_{j-1}w^{(j-1)}(x_j) + x_{j+1}w^{(j+1)}(x_j) + x_{j+2}w^{(j+2)}(x_j), \\
R_j^{[13]} = R_j^{[31]} &= x_{j-2}^2w^{(j-2)}(x_j) + x_{j-1}^2w^{(j-1)}(x_j) + x_{j+1}^2w^{(j+1)}(x_j) + x_{j+2}^2w^{(j+2)}(x_j), \\
R_j^{[22]} &= x_{j-2}^2w^{(j-2)}(x_j) + x_{j-1}^2w^{(j-1)}(x_j) + x_{j+1}^2w^{(j+1)}(x_j) + x_{j+2}^2w^{(j+2)}(x_j), \\
R_j^{[23]} = R_j^{[32]} &= x_{j-2}^3w^{(j-2)}(x_j) + x_{j-1}^3w^{(j-1)}(x_j) + x_{j+1}^3w^{(j+1)}(x_j) + x_{j+2}^3w^{(j+2)}(x_j), \\
R_j^{[33]} &= x_{j-2}^4w^{(j-2)}(x_j) + x_{j-1}^4w^{(j-1)}(x_j) + x_{j+1}^4w^{(j+1)}(x_j) + x_{j+2}^4w^{(j+2)}(x_j),
\end{aligned}$$

sowie die Vektoren $\underline{b}(x_{j+i}) = (1, x_j + i\Delta x, (x_j + i\Delta x)^2)^T$ und $\frac{d}{dx}\underline{b}(x_{j+i}) = (0, 1, 2(x_j + i\Delta x))^T$ mit den Indizes $i \in \{0, \pm 1, \pm 2\}$ für $x_j \in \{x_3, \dots, x_{n-2}\}$. Dann liefert Theorem 2.1

$$\frac{d}{dx}\mathcal{G}f(x_j) \sim \boxed{\varphi'_{j-2}(x_j)} \bullet \bullet \boxed{\varphi'_{j-1}(x_j)} \bullet \bullet \boxed{\varphi'_j(x_j)} \bullet \bullet \boxed{\varphi'_{j+1}(x_j)} \bullet \bullet \boxed{\varphi'_{j+2}(x_j)} \quad (5.12)$$

mit

$$\begin{aligned}
\varphi'_{j-2}(x_j) &= -\frac{1}{\Delta x} \frac{1}{2\mathfrak{D}_{\alpha,\beta}} \Delta x^\alpha \left(32\Delta x^\beta + 2^\alpha(3\Delta x^\beta - \Delta x^\alpha) \right), \\
\varphi'_{j-1}(x_j) &= -\frac{1}{\Delta x} \frac{1}{2\mathfrak{D}_{\alpha,\beta}} \Delta x^\alpha \left(2^{2+\alpha}\Delta x^\alpha + 2^\beta(3 \cdot 2^2 + 2^\alpha)\Delta x^\beta \right), \\
\varphi'_j(x_j) &= \frac{1}{\Delta x} \frac{3}{2\mathfrak{D}_{\alpha,\beta}} \left(2^\alpha\Delta x^{2\alpha} - 2^\beta\Delta x^{2\beta} - 3(2^\alpha - 2^\beta)\Delta x^{\alpha+\beta} \right), \\
\varphi'_{j+1}(x_j) &= \frac{1}{\Delta x} \frac{1}{2\mathfrak{D}_{\alpha,\beta}} \Delta x^\beta \left(2^{2+\beta}\Delta x^\beta + 2^\alpha(3 \cdot 2^2 + 2^\beta)\Delta x^\alpha \right), \\
\varphi'_{j+2}(x_j) &= +\frac{1}{\Delta x} \frac{1}{2\mathfrak{D}_{\alpha,\beta}} \Delta x^\beta \left(32\Delta x^\alpha + 2^\beta(3\Delta x^\alpha - \Delta x^\beta) \right),
\end{aligned}$$

wobei wir abkürzend

$$\mathfrak{D}_{\alpha,\beta} = 2^\alpha\Delta x^{2\alpha} + 2^\beta\Delta x^{2\beta} + \left(64 + 9(2^\alpha + 2^\beta) + 2^{\alpha+\beta} \right) \Delta x^{\alpha+\beta} \quad (5.13)$$

verwenden. Der Wert $\mathfrak{D}_{\alpha,\beta}$ in (5.13) ist hierbei nicht mit dem Wert $\mathfrak{D}_{\alpha,\beta}^n$ in (5.11) zu verwechseln.

Wir erkennen sofort, dass $\mathfrak{D}_{\alpha,\beta} > 0$ für alle $\alpha, \beta \in \mathbb{N}$ ist. Eine Konsistenzanalyse von (5.12) basiert auf der Taylor-Entwicklung

$$\begin{aligned}
& \underbrace{\frac{d}{dx} \mathcal{G}f(x_j)}_{\approx \frac{d}{dx} f(x_j)} - \left(\varphi'_{j-2}(x_j) f_{j-2} + \varphi'_{j-1}(x_j) f_{j-1} + \varphi'_j(x_j) f_j + \varphi'_{j+1}(x_j) f_{j+1} + \varphi'_{j+2}(x_j) f_{j+2} \right) = \\
& \frac{\Delta x^2}{3!} \frac{1}{\mathfrak{D}_{\alpha,\beta}} \left(2^{1+\alpha} \Delta x^{2\alpha} + 2^{1+\beta} \Delta x^{2\beta} - (256 + 9(2^{1+\alpha} + 2^{1+\beta}) + 2^{\alpha+\beta}) \Delta x^{\alpha+\beta} \right) \frac{d^3}{dx^3} f_j \\
& - \frac{\Delta x^3}{4} \frac{1}{\mathfrak{D}_{\alpha,\beta}} \left(2^\alpha \Delta x^{2\alpha} - 2^\beta \Delta x^{2\beta} - 3(2^\alpha - 2^\beta) \Delta x^{\alpha+\beta} \right) \frac{d^4}{dx^4} f_j \\
& + \frac{\Delta x^4}{5!} \frac{1}{\mathfrak{D}_{\alpha,\beta}} \left(7(2^{1+\alpha} \Delta x^{2\alpha} + 2^{1+\beta} \Delta x^{2\beta}) - (1024 + 27(2^{1+\alpha} + 2^{1+\beta}) + 2^{\alpha+\beta}) \Delta x^{\alpha+\beta} \right) \frac{d^5}{dx^5} f_j \\
& - \frac{\Delta x^5}{24} \frac{1}{\mathfrak{D}_{\alpha,\beta}} \left(2^\alpha \Delta x^{2\alpha} - 2^\beta \Delta x^{2\beta} - 3(2^\alpha - 2^\beta) \Delta x^{\alpha+\beta} \right) \frac{d^6}{dx^6} f_j \\
& + \frac{\Delta x^6}{7!} \frac{1}{\mathfrak{D}_{\alpha,\beta}} \left(31(2^{1+\alpha} \Delta x^{2\alpha} + 2^{1+\beta} \Delta x^{2\beta}) - (4096 + 99(2^{1+\alpha} + 2^{1+\beta}) + 2^{\alpha+\beta}) \Delta x^{\alpha+\beta} \right) \frac{d^7}{dx^7} f_j \\
& + \mathcal{O}(\Delta x^7),
\end{aligned}$$

mit $\mathfrak{D}_{\alpha,\beta}$ wie in Gleichung (5.13). Daher ist die Finite Differenz (5.12) konsistent von zweiter Ordnung für jede Wahl von α und β , für die

$$2^{1+\alpha} \Delta x^{2\alpha} + 2^{1+\beta} \Delta x^{2\beta} - (256 + 9(2^{1+\alpha} + 2^{1+\beta}) + 2^{\alpha+\beta}) \Delta x^{\alpha+\beta} \neq 0 \quad (5.14)$$

erfüllt ist. Insbesondere gilt (5.14) für den Fall $\alpha = \beta$, denn die Gleichung

$$\begin{aligned}
2^{2+\alpha} \Delta x^{2\alpha} - (256 + 9 \cdot 2^{2+\alpha} + 2^{2\alpha}) \Delta x^{2\alpha} &= 0 & \iff \\
2^{2+\alpha} &= 256 + 9 \cdot 2^{2+\alpha} + 2^{2\alpha} & \iff \\
0 &= 256 + 8 \cdot 2^{2+\alpha} + 2^{2\alpha} & \iff \\
0 &= 256 + 32 \cdot 2^\alpha + (2^\alpha)^2 & \iff \\
0 &= (16 + 2^\alpha)^2
\end{aligned}$$

besitzt keine Lösung $\alpha \in \mathbb{R}$.

Aus (5.9) bzw. (5.12) ein Verfahren zur numerischen Approximation der nichtlinearen Advektionsgleichung durch

$$U_j^{n+1} = U_j^n - \Delta t \frac{d}{dx} \mathcal{G}f(U_j^n),$$

das heißt

$$U_j^{n+1} = U_j^n - \Delta t \sum_{k=j-2}^{j+2} \varphi'_k(x_j) f(U_k^n), \quad (5.15)$$

mit den Werten $\varphi'_i(x_j)$ für $j-2 \leq i \leq j+2$ entsprechend denen in (5.9) bzw. (5.12).

Satz 5.2 *Durch (5.15) ist ein konservatives und konsistentes Verfahren zur numerischen Behandlung der nichtlinearen Advektionsgleichung (5.1a) gegeben, falls die Werte $\phi'_k(x_j)$ für die Indizes $j-2 \leq k \leq j+2$ gemäß (5.9) bzw. (5.12) berechnet werden.*

Beweis. Zunächst zeigen wir die Konservativität der Methode mit den Werten $\phi'_k(x_j)$ aus (5.9). Mit Hilfe der numerischen Flussfunktion

$$\begin{aligned} F_{\text{lin}}(v_1, v_2, v_3, v_4) &= \frac{2^{1+\alpha}\Delta x^\alpha}{\mathfrak{D}_{\alpha,\beta}^2} f(v_1) + \frac{2^\alpha(2+2^\beta)\Delta x^\alpha}{\mathfrak{D}_{\alpha,\beta}^2} f(v_2) \\ &\quad + \frac{2^\beta(2+2^\alpha)\Delta x^\beta}{\mathfrak{D}_{\alpha,\beta}^2} f(v_3) + \frac{2^{1+\beta}\Delta x^\beta}{\mathfrak{D}_{\alpha,\beta}^2} f(v_4), \end{aligned}$$

für $\mathfrak{D}_{\alpha,\beta}^2$ gemäß (5.11), kann der Ausdruck (5.15) in der Form

$$U_j^{n+1} = U_j^n - \frac{\Delta t}{\Delta x} \left(F_{\text{lin}}(U_{j-1}^n, U_j^n, U_{j+1}^n, U_{j+2}^n) - F_{\text{lin}}(U_{j-2}^n, U_{j-1}^n, U_j^n, U_{j+1}^n) \right)$$

geschrieben werden. Weiter gilt für den Fall $\phi'_k(x_j)$ entsprechend (5.12), dass vermöge der numerischen Flussfunktion

$$\begin{aligned} F_{\text{quad}}(v_1, v_2, v_3, v_4) &= -\frac{1}{2\mathfrak{D}_{\alpha,\beta}} \Delta x^\alpha \left(2^\alpha \Delta x^\alpha - 32\Delta x^\beta - 3 \cdot 2^\alpha \Delta x^\beta \right) f(v_1) \\ &\quad + \frac{1}{2\mathfrak{D}_{\alpha,\beta}} \Delta x^\alpha \left(3 \cdot 2^\alpha \Delta x^\alpha + 32\Delta x^\beta + 3 \cdot 2^\alpha \Delta x^\beta + 3 \cdot 2^{2+\beta} \Delta x^\beta + 2^{\alpha+\beta} \Delta x^\beta \right) f(v_2) \\ &\quad + \frac{1}{2\mathfrak{D}_{\alpha,\beta}} \Delta x^\beta \left(32\Delta x^\alpha + 3 \cdot 2^{2+\alpha} \Delta x^\alpha + 3 \cdot 2^\beta \Delta x^\alpha + 2^{\alpha+\beta} \Delta x^\alpha + 3 \cdot 2^\beta \Delta x^\beta \right) f(v_3) \\ &\quad - \frac{1}{2\mathfrak{D}_{\alpha,\beta}} \Delta x^\beta \left(-32\Delta x^\alpha - 3 \cdot 2^\beta \Delta x^\alpha + 2^\beta \Delta x^\beta \right) f(v_4) \end{aligned}$$

der Ausdruck (5.15) in der Form

$$U_j^{n+1} = U_j^n - \frac{\Delta t}{\Delta x} \left(F_{\text{quad}}(U_{j-1}^n, U_j^n, U_{j+1}^n, U_{j+2}^n) - F_{\text{quad}}(U_{j-2}^n, U_{j-1}^n, U_j^n, U_{j+1}^n) \right)$$

geschrieben werden kann. An dieser Stelle verzichten wir auf eine Rechnung analog zu (5.8) und notieren, dass ein direktes Nachrechnen unmittelbar die Konsistenzbedingungen

$$F_{\text{lin}}(\tilde{u}, \tilde{u}, \tilde{u}, \tilde{u}) = f(\tilde{u})$$

und

$$F_{\text{quad}}(\tilde{u}, \tilde{u}, \tilde{u}, \tilde{u}) = f(\tilde{u})$$

für alle $\tilde{u} \in \mathbb{R}$ liefert. □

5.3 Ein informeller Exkurs über Upwind-Verfahren

Betrachten wir noch einmal das Cauchy-Problem für die nichtlineare Advektionsgleichung

$$\begin{aligned} u(x, t)_t + f(u(x, t))_x &= 0, & x \in \mathbb{R}, \quad t > 0, \\ u(x, 0) &= u_0(x), & x \in \mathbb{R}. \end{aligned}$$

Für den Fall einer linearen Flussfunktion f , das heißt $f(u) = au$ für $a \in \mathbb{R}$, existiert zu gegebenem Anfangsdatum $u_0 \in C^1(\mathbb{R}; \mathbb{R})$ für alle Zeiten $t > 0$ eine eindeutig bestimmte klassische Lösung

$$u(x, t) = u_0(x - at). \quad (5.16)$$

Dies können wir zum einen durch direktes Nachrechnen verifizieren und zum anderen mit Hilfe des Begriffes der *Charakteristiken* der Advektionsgleichung erklären: Jede Lösung der linearen Gleichung $u_t + au_x = 0$ ist konstant auf den Kurven

$$x - at = \text{const} \quad (5.17)$$

in der x - t -Ebene. Denn ist $x = x(t)$, gilt aufgrund von $\frac{dx(t)}{dt} = \frac{d}{dt}(at + \text{const}) = a$ für jede differenzierbare Lösung u von $u_t + au_x = 0$ schon

$$\frac{d}{dt}u(x(t), t) = u_t + u_x \frac{dx(t)}{dt} = u_t + au_x = 0. \quad (5.18)$$

Die Kurven (5.17) werden als *Charakteristiken* bezeichnet. Die Charakteristiken der Gleichung $u_t + au_x = 0$ können durch Integration der gewöhnlichen Differentialgleichung

$$\frac{dx(t)}{dt} = a$$

bestimmt werden. Allgemein sind Charakteristiken für die nichtlineare Advektionsgleichung $u_t + f(u)_x = 0$ durch die gewöhnliche Differentialgleichung

$$\frac{dx(t)}{dt} = f'(u(x(t), t)) \quad (5.19)$$

bestimmt, denn ist u eine Lösung von $u_t + f(u)_x = 0$, so gilt

$$\frac{d}{dt}u(x(t), t) = u_t + u_x \frac{dx(t)}{dt} = u_t + u_x f'(u(x(t), t)) = u_t + f(u)_x = 0.$$

Das einfachste Exemplar einer nichtlinearen Advektionsgleichung ist die so genannte *Burgers-Gleichung*

$$u_t + \left(\frac{1}{2}u^2\right)_x = 0. \quad (5.20)$$

Hierbei ist $f(u) = \frac{1}{2}u^2$ und $f'(u) = u$. Mit Blick auf (5.19) sind die Charakteristiken der Burgers-Gleichung Geraden in der x - t -Ebene, parametrisiert in der Form

$$x(t) = x_0 + u_0(x_0)t, \quad (5.21)$$

wobei (5.20) mit dem Anfangsdatum u_0 ausgestattet sei. Aufgrund von (5.19) und (5.21) gilt

$$\begin{aligned} \frac{d}{dt}u(x(t), t) &= \frac{d}{dt}u(x_0 + u_0(x_0)t, t) = u_t + u_x u_0(x_0) \\ &= u_t + u_x \frac{dx(t)}{dt} = u_t + u_x u = u_t + \left(\frac{1}{2}u^2\right)_x = 0, \quad x(0) = x_0. \end{aligned}$$

Werfen wir an dieser Stelle einen Blick auf eine typische Charakteristiken-Konfiguration für die lineare Advektionsgleichung, die in Abbildung 5.1 dargestellt ist.

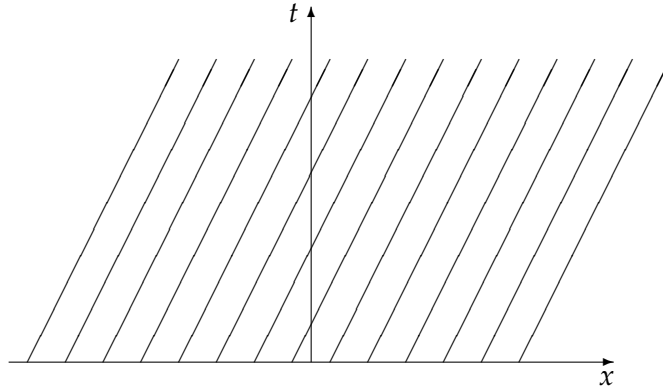


Abbildung 5.1: Charakteristiken-Konfiguration für die lineare Advektionsgleichung.

Die Strategie zur Lösung des Cauchy-Problems der linearen Advektionsgleichung stellt sich mit Blick auf Abbildung 5.1 wie folgt dar: Der Wert $u(x, t)$ der Lösung im Punkt (x, t) ist dadurch zu berechnen, indem die durch diesen Punkt verlaufende Charakteristik in der x - t -Ebene bis zur x -Achse ($t = 0$) zurückverfolgt wird. Die entsprechende Charakteristik schneide die x -Achse im Punkt $(\tilde{x}, 0)$. Dann ist $u(\tilde{x}, 0)$ durch das Anfangsdatum bestimmt, das heißt $u(\tilde{x}, 0) = u_0(\tilde{x})$ und aufgrund von (5.18) gilt $u(x, t) = u(\tilde{x}, 0)$. Diese Lösungsstrategie liefert für alle Zeiten $t > 0$ eindeutig bestimmte Lösungswerte des Cauchy-Problems.

Beim Versuch diese Strategie auf das Cauchy-Problem für die Burgers-Gleichung zu übertragen, stellen wir fest, dass sich jeweils zwei Charakteristiken

$$x(t) = x_0 + u_0(x_0)t \quad \text{und} \quad x(t) = x_1 + u_0(x_1)t$$

mit $u_0(x_0) \neq u_0(x_1)$ zum Zeitpunkt

$$t = \frac{x_0 - x_1}{u_0(x_1) - u_0(x_0)}$$

schneiden. Das heißt, bei $t = T$ mit

$$T = \inf_{x_0, x_1 \in \mathbb{R}} \frac{x_0 - x_1}{u_0(x_1) - u_0(x_0)} = \frac{-1}{\inf_{x \in \mathbb{R}} u'_0(x)}$$

schneiden sich zwei verschiedene Charakteristiken der Burgers-Gleichung. Sei beispielsweise (\tilde{x}, \tilde{t}) dieser Schnittpunkt. Nehmen wir $T > 0$ an, so bricht unsere klassische Lösungsstrategie zusammen, denn durch (\tilde{x}, \tilde{t}) laufen mindestens zwei Charakteristiken, und der Wert der Lösung

an dieser Stelle ist nicht mehr eindeutig bestimmt. Genauer besitzt das Cauchy-Problem für die Burgers-Gleichung für $t \geq T$ keine differenzierbare Lösung. Durch die Begriffsbildung der schwachen bzw. Entropie-Lösung lässt sich diese Problematik umgehen.

Wenden wir uns unabhängig von diesen verallgemeinerten Lösungsbegriffen noch einmal der Abbildung 5.1 zu. Legen wir hierbei über die x - t -Ebene das Gitter

$$\left\{ (j \Delta x, n \Delta t) \mid j \in \mathbb{Z}, n \in \mathbb{N}_0 \right\},$$

so resultiert:

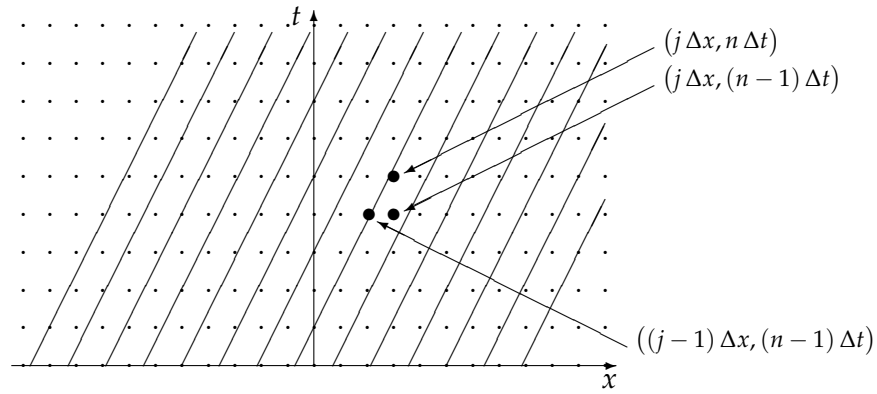


Abbildung 5.2: x - t -Ebene mit Gitter.

Angenommen, die Lösung des Cauchy-Problems für die lineare Advektionsgleichung ist zum Zeitpunkt $(n-1)\Delta t$ für alle Werte $j\Delta x$ bekannt bzw. durch ein numerisches Verfahren approximiert. Diese Werte seien durch die Menge $\{U_j^{n-1}\}_{j \in \mathbb{N}}$ gegeben. Dann stellt sich die Frage, wie der Wert U_j^n der Lösung am Punkt $(j\Delta x, n\Delta t)$ mit Blick auf die reale Transportrichtung des Anfangsdatums entlang der Charakteristiken „sinnvoll“ zu approximieren ist. Es ist ein klassisches Resultat [18], dass von den drei Ansätzen

$$U_j^n = U_j^n(U_j^{n-1}, U_{j-1}^{n-1}), \quad (5.22)$$

$$U_j^n = U_j^n(U_{j+1}^{n-1}, U_j^{n-1}), \quad (5.23)$$

$$U_j^n = U_j^n(U_{j+1}^{n-1}, U_{j-1}^{n-1}), \quad (5.24)$$

für die in Abbildung 5.2 dargestellte Charakteristiken-Konfiguration, nur der Ansatz (5.22) die gesuchte „sinnvolle“ Approximation für den Wert U_j^n liefern kann. „Sinnvoll“ meint, dass wir uns nun streng genommen mit den Begriffen *lineare Stabilität* und *Courant-Friedrichs-Levy-Bedingung* auseinandersetzen müssten.

Wir verzichten auf diese rigorosen Begriffsbildungen und geben stattdessen ein heuristisches Argument für die Wahl von (5.22): Würden wir in einer Konfiguration wie in Abbildung 5.2 dargestellt versuchen, den Wert der Lösung bei $(j\Delta x, n\Delta t)$ unter Zuhilfenahme des Wertes im

Punkt $((j+1)\Delta x, (n-1)\Delta t)$ zu approximieren, verknüpfen wir Werte des transportierten Anfangsdatums, die in keinerlei Zusammenhang stehen. Dieses Verhalten zeigt Abbildung 5.3.

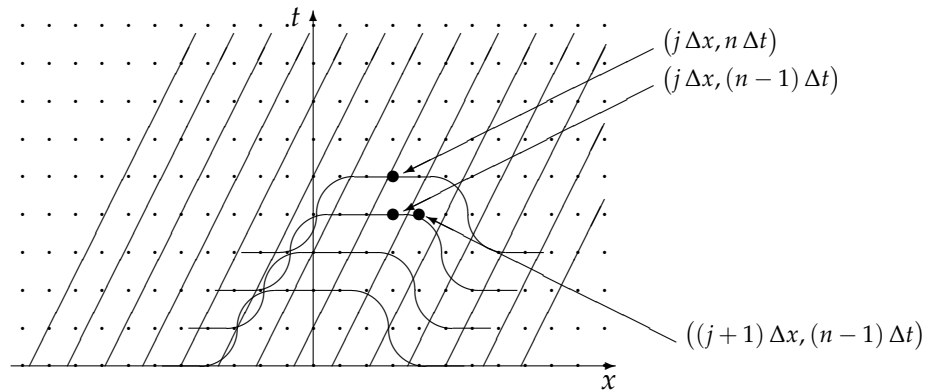


Abbildung 5.3: Transport des Anfangsdatums entlang der Charakteristiken.
Approximation mit Ansatz (5.23).

Aus dieser Betrachtung heraus erkennen wir, dass auch (5.24) den Wert U_j^n mit dem hierfür nicht relevanten U_{j+1}^{n-1} approximiert. Hingegen liefert (5.22) mit Blick auf Abbildung 5.4 für die gegebene Charakteristiken-Konfiguration, und damit der Transportrichtung des Anfangsdatums in der x - t -Ebene, eine sinnvolle Verknüpfung der Lösungswerte U_j^n und U_{j-1}^{n-1} .

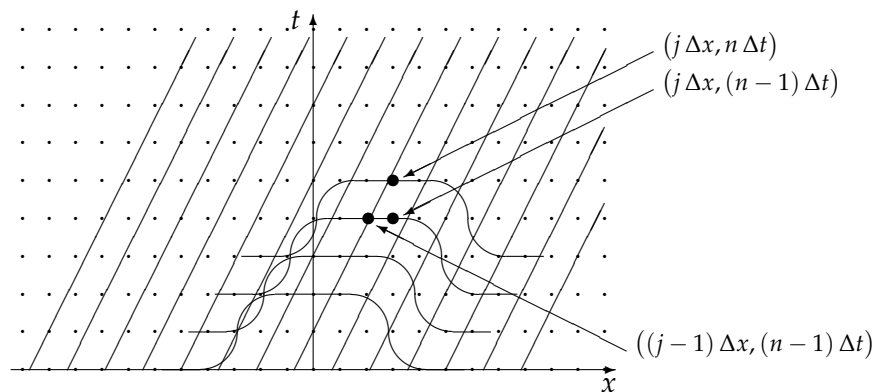


Abbildung 5.4: Transport des Anfangsdatums entlang der Charakteristiken.
Approximation mit Ansatz (5.22).

Die gegebenen Erörterungen legen es nahe, dass bei der Verwendung von Finite Differenzen für die Approximation der räumlichen Ableitung der linearen Advektionsgleichung auf eine *einseitige Ausrichtung* der entsprechenden Differenzen mit Blick auf die tatsächliche Transportrichtung zu achten ist. Gleiches gilt auch insbesondere bei der Approximation von unstetigen Lösungen der Burgers-Gleichung bzw. allgemein bei der numerischen Approximation von nicht-linearen Advektionsgleichungen durch Finite Differenzen-Verfahren.

Verfahren die diese natürliche Transportrichtung respektieren bzw. berücksichtigen, werden als *Upwind-Verfahren* bezeichnet.

5.4 Ad hoc Upwind-Verfahren durch Variation der Parameter des Interpolationsgewichtes

In Abschnitt 2.4 hatten wir bereits festgestellt, dass sich die Approximation (2.14) aus einer Vorwärts-, Rückwärts-, sowie zentralen Finite Differenz zusammensetzt, die basierend auf der Wahl der vorhandenen Parameter des unterliegenden Interpolationsgewichtes entsprechend gewichtet werden können. Diese Möglichkeit der Gewichtung liefert mit Blick auf Abschnitt 5.3 eine ad hoc-Methode zur Konstruktion von im Wesentlichen einseitigen Finite Differenzen. Ganz allgemein gilt für die Ableitungen der Kernfunktionen nach Theorem 2.1

$$D^\mu \varphi_i(x_j) = w^{(i)}(x_j) \underline{b}(x_i)^T \left(R(x_j)^{-1} - \frac{1}{\sigma(x_j, x_j)} (R(x_j)^{-1} \underline{b}(x_j)) (R(x_j)^{-1} \underline{b}(x_j))^T \right) D^\mu \underline{b}(x_j).$$

Das heißt, die Ableitungen sind von der Struktur

$$D^\mu \varphi_i(x_j) = w(x_j - x_i) \cdot (\dots\dots\dots). \quad (5.25)$$

Nehmen wir für w in (5.25) an, dass es ein unsymmetrisches und ungewichtetes Norm-inverses positives Interpolationsgewicht nach Definition 2.2 ist, so resultiert

$$D^\mu \varphi_i(x_j) = \begin{cases} \frac{1}{|x_j - x_i|^\alpha} \cdot (\dots\dots\dots) & : x_i > x_j \\ \frac{1}{|x_j - x_i|^\beta} \cdot (\dots\dots\dots) & : x_i < x_j \end{cases}, \quad (5.26)$$

und damit eine Gewichtungsmöglichkeit der hieraus zu bestimmenden Finite Differenz durch Variation der Parameter α und β . Analoge Betrachtungen gelten auch für Träger-eingeschränkte positive Interpolationsgewichte.

Betrachten wir anhand einer Verallgemeinerung von (5.15), das heißt anhand der Methode

$$U_j^{n+1} = U_j^n - \Delta t \sum_{k=j-s}^{j+s} \varphi'_k(x_j) f(U_k^n), \quad s \in \mathbb{N}, \quad (5.27)$$

wie ein ad hoc Upwind-Verfahren für die nichtlineare Advektionsgleichung konstruiert werden kann: Die Approximation der räumlichen Ableitung der nichtlinearen Advektionsgleichung ist in (5.27) durch den Term

$$\sum_{k=j-s}^{j+s} \varphi'_k(x_j) f(U_k^n) \quad (5.28)$$

gegeben, das heißt durch den Wert der Ableitung einer entsprechenden exakte kleinste Quadrate Interpolierenden bei $x = x_j$. Die Variation der Parameter des Interpolationsgewichtes impliziert dann beispielsweise aufgrund von (5.26) die entsprechende Ausrichtung und Gewichtung der Faktoren in (5.28), falls in diesem Fall wieder die Verwendung eines ungewichteten und unsymmetrischen Norm-inversen positiven Interpolationsgewichtes angenommen wird.

5.5 Numerische Beispiele

Im Folgenden testen wir die Methode (5.27) für $s = 1$ und $s = 2$ anhand eines Cauchy-Problems für die lineare Advektionsgleichung

$$u_t + u_x = 0 \quad (5.29)$$

bzw. die Burgers-Gleichung

$$u_t + \left(\frac{1}{2}u^2\right)_x = 0. \quad (5.30)$$

Dazu stellen wir die Gleichungen jeweils mit dem stückweise differenzierbaren Anfangsdatum

$$u_0(x) = \begin{cases} 1 & : x \leq 0 \\ 1 - x & : 0 \leq x \leq 1 \\ 0 & : 1 \leq x \end{cases} \quad (5.31)$$

aus. Die Entropie-Lösung von (5.29), (5.31) ist

$$u(x, t) = \begin{cases} 1 & : x \leq t \\ 1 - (x - t) & : t \leq x \leq t + 1 \\ 0 & : t + 1 \leq x \end{cases},$$

und für (5.30), (5.31) erhalten wir für Zeiten $t < 1$ die Entropie-Lösung

$$u(x, t) = \begin{cases} 1 & : x \leq t \\ (1 - x)/(1 - t) & : t \leq x \leq 1 \\ 0 & : x \geq 1 \end{cases},$$

sowie $u(x, t) = \begin{cases} 1 & : x < (t + 1)/2 \\ 0 & : x > (t + 1)/2 \end{cases}$ für alle Zeiten $t \geq 1$, mit einem Verweis auf [7].

Zunächst widmen wir uns der numerischen Behandlung des Cauchy-Problems für die lineare Advektionsgleichung (5.29) für $x \in [-1, 3]$. Wir wählen von nun an die Parameter $\Delta x = \frac{1}{20}$, $\alpha = 1$ und $\beta = 3$ fest für alle folgenden Beispiele.

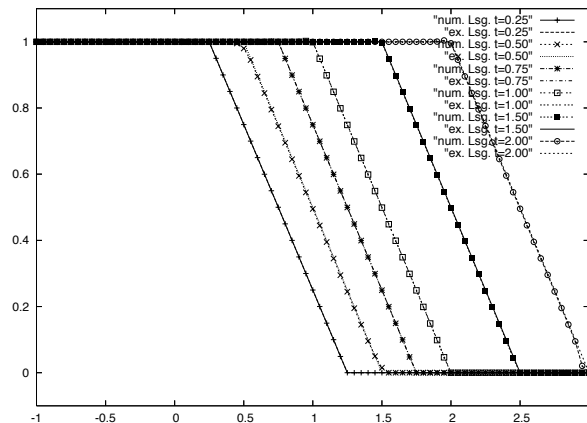


Abbildung 5.5: Numerische Approximation des Cauchy-Problems für (5.29) aus (5.27) mit einem 3-Punkte Upwind-Verfahren unter Verwendung von (2.14).

Die Methode (5.27) mit $s = 1$ und den $\varphi'_k(x_j)$ entsprechend denen in (2.14) ausgestattet, liefert für Zeiten $t \in \{0.25, 0.5, 0.75, 1.0, 1.5, 2.0\}$ die Ergebnisse in Abbildung 5.5. Diese und auch die folgenden Abbildungen zeigen jeweils die exakte (Entropie-) Lösung («ex. Lsg. $t=x.yz$ ») und die numerische Lösung («num. Lsg. $t=x.yz$ »). Verwenden wir die Werte $\varphi'_k(x_j)$ aus (5.9), das heißt wir nutzen (5.27) mit $s = 2$, so resultieren die Ergebnisse wie in Abbildung 5.6 dargestellt.

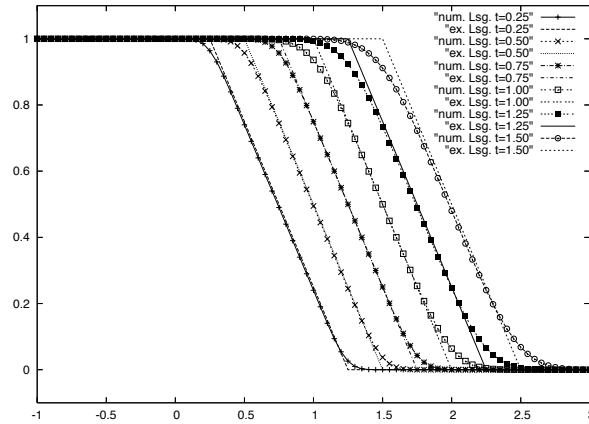


Abbildung 5.6: Numerische Approximation des Cauchy-Problems für (5.29) mit einem 5-Punkte Upwind-Verfahren aus (5.27) unter Verwendung von (5.9).

Die bisher genutzten Verfahren sind jeweils Verfahren von erster Ordnung bezüglich der zeitlichen und räumlichen Ableitungs-Approximationen. Die Wahl der $\varphi'_k(x_j)$ aus (5.12) liefert ein Verfahren mit räumlichen Ableitungs-Approximationen von zweiter Ordnung. In Abbildung 5.7 sind die Ergebnisse für $t \in \{0.25, 0.5\}$ dargestellt.

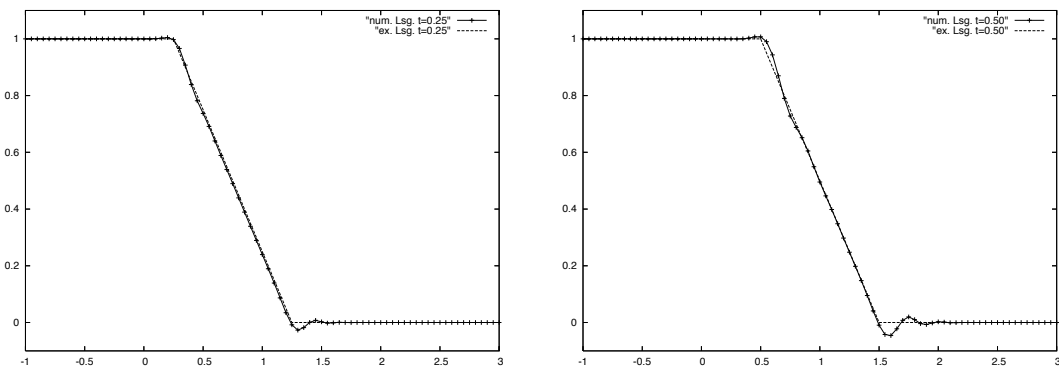


Abbildung 5.7: 5-Punkte Upwind-Verfahren des Cauchy-Problems für (5.29) aus (5.27) mit räumlicher Ableitungs-Approximation von zweiter Ordnung für $t = 0.25$ (links) und $t = 0.5$ (rechts) unter Verwendung von (5.12).

Wir erkennen hierbei aufkommende Oszillationen, die sich für größer werdende Zeiten t weiter verstärken. Fassen wir an dieser Stelle kurz die bisherigen Ergebnisse zusammen: Nutzen wir ein 3-Punkte Upwind-Verfahren auf Basis von (5.27) wie in Abbildung 5.5 gezeigt, sind die numerisch berechneten Lösungen für das gegebene Cauchy-Problem nahezu identisch zu den

exakten Lösungen. Erweitern wir hingegen auf eine 5-Punkte Finite Differenz für die Approximation der räumlichen Ableitung (mit einer Genauigkeit von erster Ordnung) wie in Abbildung 5.6 dargestellt, verschmiert die numerische Lösung an Punkten, die zu den Punkten der exakten Lösung korrespondieren, an denen diese nicht differenzierbar ist.

Sämtliche Rechnungen führen wir im Folgenden noch einmal unter analogen Voraussetzungen für das entsprechende Cauchy-Problem für die Burgers-Gleichung (5.30) durch.

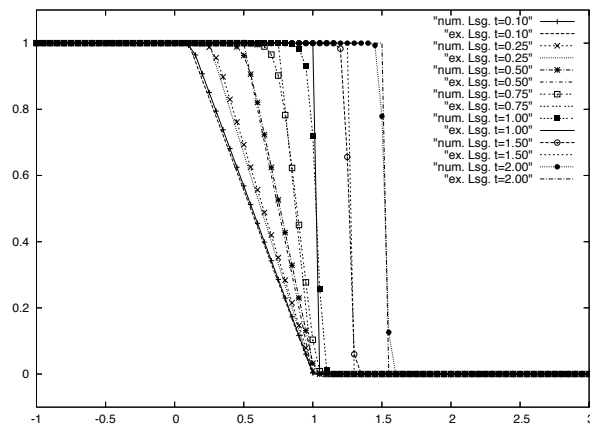


Abbildung 5.8: Numerische Approximation des Cauchy-Problems für (5.30) mit einem 3-Punkte Upwind-Verfahren aus (5.27) unter Verwendung von (2.14).

Die Abbildung 5.8 zeigt die entsprechenden Ergebnisse für ein 3-Punkte Upwind-Verfahren unter Verwendung von (2.14). Für ein 5-Punkte Upwind-Verfahren unter Verwendung von (5.9) resultieren die Ergebnisse in Abbildung 5.9.

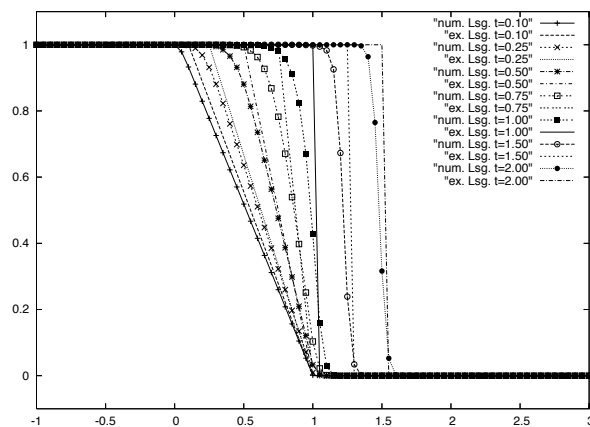


Abbildung 5.9: Numerische Approximation des Cauchy-Problems für (5.30) mit einem 5-Punkte Upwind-Verfahren aus (5.27) unter Verwendung von (5.9).

Ähnlich zum Fall der linearen Advektionsgleichung sehen wir in Abbildung 5.9 ein Verschmieren der numerischen Lösung und stellen weiter fest, dass die Ergebnisse des Verfahrens aus Abbildung 5.8 verglichen mit Abbildung 5.5 wesentlich schlechtere Approximationen an die exakte

Lösung liefert. Ferner entstehen aus einem 5-Punkte Upwind-Verfahren auf Basis von (5.27), das mit einer räumlichen Ableitungs-Approximation von zweiter Ordnung ausgestattet ist, Oszillationen um Punkte in der numerischen Lösung, die zu Punkten der exakten Lösung korrespondieren, an denen diese nicht differenzierbar ist. Die Resultate sind in Abbildung 5.10 dargestellt.

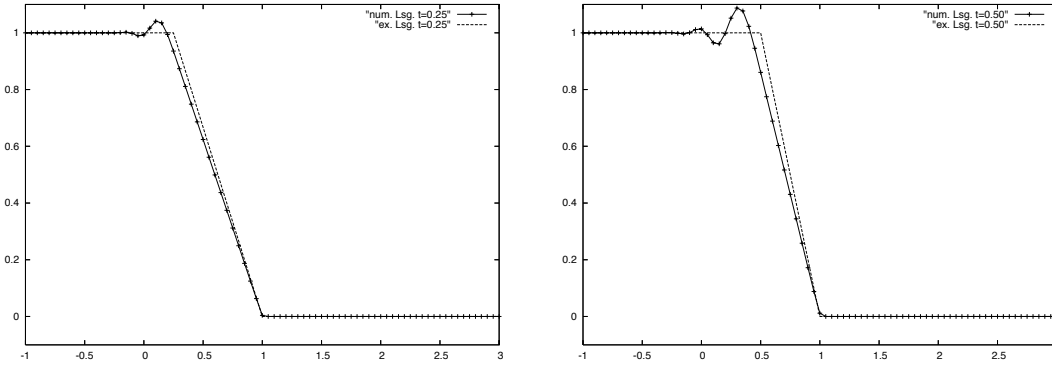


Abbildung 5.10: 5-Punkte Upwind-Verfahren für (5.30) aus (5.27) mit räumlicher Ableitungs-Approximation von zweiter Ordnung für $t = 0.25$ (links) und $t = 0.5$ (rechts) unter Verwendung von (5.12).

Zur Dämpfung der gerade beschriebenen Oszillationen etablieren wir im Folgenden ein numerisches Verfahren zur Lösung der nichtlinearen Advektionsgleichung vom *Lax-Wendroff*-Typ. Hierfür sei angenommen, dass die Funktion u eine differenzierbare Lösung der Gleichung $u_t + f(u)_x = 0$ ist. Dann liefert eine Taylor-Entwicklung für kleine Zeitschrittweiten Δt

$$u(x, t + \Delta t) = u(x, t) + \Delta t u_t(x, t) + \frac{(\Delta t)^2}{2} u_{tt}(x, t) + \mathcal{O}((\Delta t)^3). \quad (5.32)$$

Aus $u_t = -f(u)_x$ und $u_{tt} = -f(u)_{xt} = -f(u)_{tx} = -(f'(u) u_t)_x = (f'(u) f(u)_x)_x$ folgt dann schon mit Blick auf (5.32)

$$u(x, t + \Delta t) = u(x, t) - \Delta t f(u(x, t))_x + \frac{(\Delta t)^2}{2} (f'(u(x, t)) f(u(x, t))_x)_x. \quad (5.33)$$

Aus (5.33) erhalten wir in Analogie zu (5.15) durch

$$U_j^{n+1} = U_j^n - \Delta t \frac{d}{dx} \mathcal{G} f(U_j^n) + \frac{(\Delta t)^2}{2} \frac{d}{dx} \mathcal{G} \left(f'(U_j^n) \frac{d}{dx} \mathcal{G} f(U_j^n) \right) \quad (5.34)$$

eine Methode zur numerischen Approximation der nichtlinearen Advektionsgleichung, wobei \mathcal{G} den Operator bezüglich einer exakte kleinste Quadrate Interpolation bezeichnet. Die Methode (5.34) ist hierbei ein Verfahren von zweiter Ordnung bezüglich der Ableitungs-Approximationen in der Zeit.

Zunächst widmen wir uns der numerischen Behandlung des Cauchy-Problems für die lineare Advektionsgleichung (5.29). Aus (5.34) resultiert hierfür die Methode

$$U_n^{n+1} = U_j^n - \Delta t \frac{d}{dx} \mathcal{G} f(U_j^n) + \frac{(\Delta t)^2}{2} \frac{d}{dx} \mathcal{G} \left(\frac{d}{dx} \mathcal{G} f(U_j^n) \right). \quad (5.35)$$

Wir nehmen im Folgenden wieder $x \in [-1, 3]$, $\alpha = 1$ sowie $\beta = 3$ an und verzichten auf Rechnungen mit 3-Punkte Verfahren. Dann liefert die Methode (5.35) unter Anwendung der Finite Differenz (5.9) die Ergebnisse wie in Abbildung 5.11 dargestellt.

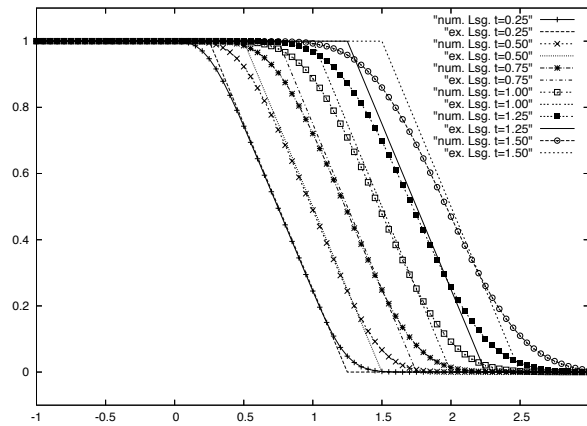


Abbildung 5.11: Numerische Approximation des Cauchy-Problems für (5.29) mit einem 5-Punkte Verfahren vom Lax-Wendroff-Typ (5.35) unter Verwendung von (5.9).

Nutzen wir hingegen die Finite Differenz (5.12), wird (5.35) zu einem Verfahren von zweiter Ordnung sowohl in der Zeit als auch im Raum. Die entsprechenden Resultate des Verfahrens zeigt die Abbildung 5.12.

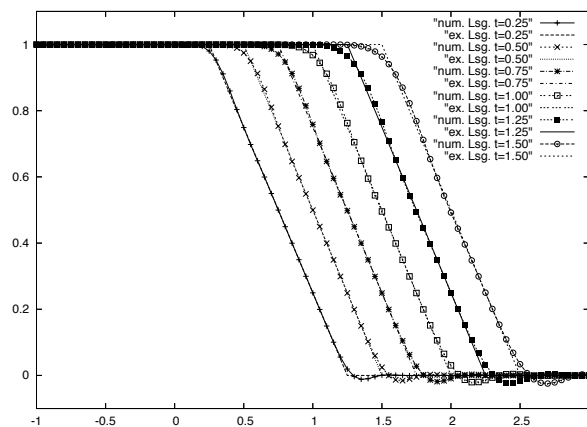


Abbildung 5.12: Numerische Approximation des Cauchy-Problems für (5.29) mit einem 5-Punkte Verfahren vom Lax-Wendroff-Typ (5.35) unter Verwendung von (5.12).

Dieses Verfahren rein von zweiter Ordnung liefert im Vergleich zu dem mit einer räumlichen Ableitungs-Approximation von erster Ordnung wie in Abbildung 5.11 weniger verschmierte numerische Lösungen, jedoch erkennen wir ein leichtes Aufkommen von Oszillationen, die sich allerdings für größer werdende Zeiten nicht verstärken.

Für die Burgers-Gleichung mit $f(u) = \frac{1}{2}u$ erhalten wir in Analogie zu (5.35) das Verfahren

$$U_n^{n+1} = U_j^n - \Delta t \frac{d}{dx} \mathcal{G} f(U_j^n) + \frac{(\Delta t)^2}{2} \frac{d}{dx} \mathcal{G} \left(U_j^n \frac{d}{dx} \mathcal{G} \left(\frac{1}{2} (U_j^n)^2 \right) \right). \quad (5.36)$$

Wenden wir dieses Verfahren zunächst mit (5.9) an, das heißt das Verfahren ist von erster Ordnung bezüglich der räumlichen Ableitungs-Approximation, so resultieren die Ergebnisse in Abbildung 5.12. Wieder kann ein Verschmieren der numerischen Lösung insbesondere für größere Zeiten um die entstehende Unstetigkeit der exakten Lösung herum festgestellt werden.

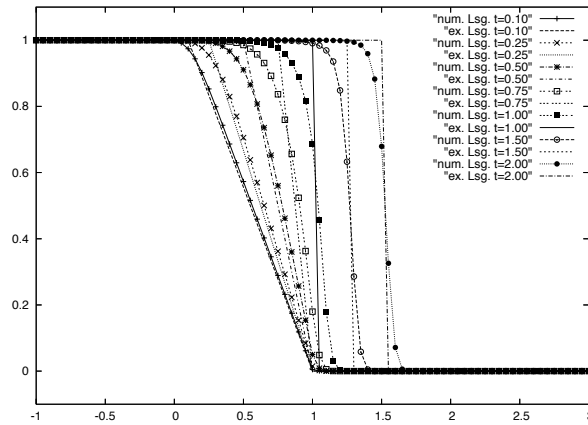


Abbildung 5.12: Numerische Approximation des Cauchy-Problems für (5.30) mit einem 5-Punkte Verfahren vom Lax-Wendroff-Typ unter Verwendung von (5.9).

Ein Verfahren rein von zweiter Ordnung erhalten wir aus (5.36) unter Verwendung von (5.12). Die entsprechenden Ergebnisse sind in Abbildung 5.13 dargestellt.

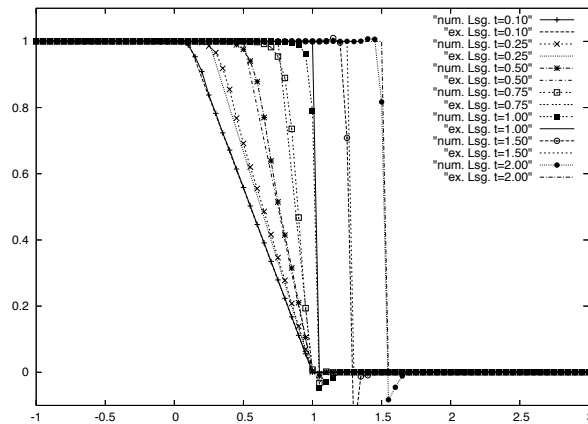


Abbildung 5.13: Numerische Approximation des Cauchy-Problems für (5.30) mit einem 5-Punkte Verfahren vom Lax-Wendroff-Typ (5.36) unter Verwendung von (5.12).

Hierbei erkennen wir aufkommende Oszillationen in Bereichen der numerischen Lösung, in denen die korrespondierende exakte Lösung eine Unstetigkeit aufweist. Für kleine Zeitbereiche

bleibt der Typ dieser Oszillation zunächst erhalten und verstärkt sich nicht. Schauen wir uns daher diesen Oszillations-Typ näher an: In Abbildung 5.14 ist für den Zeitpunkt $t = 2.0$ die numerische Lösung und exakte Lösung des Cauchy-Problems für (5.30) dargestellt.

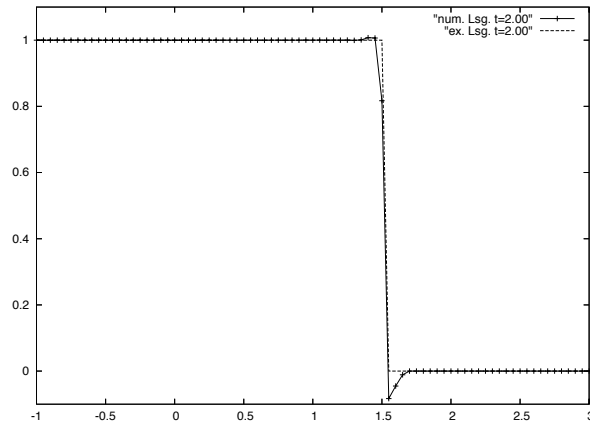


Abbildung 5.14: Numerische Approximation des Cauchy-Problems für (5.30) mit einem 5-Punkte Verfahren vom Lax-Wendroff-Typ (5.36) unter Verwendung von (5.12) für $t = 2.0$

Die Unstetigkeit der exakten Lösung wird im Bereich links von dieser durch die numerische Lösung gut approximiert. Die erwähnten Oszillationen treten nur in einem kleinen Bereich rechts der Unstetigkeit auf. Wir nehmen daher an, dass dieses Verhalten zum einen auf der einseitigen Gewichtung der zur Approximation der räumlichen Ableitung verwendeten Finite Differenz und zum anderen auf deren Genauigkeit von zweiter Ordnung beruht. Unter diesen Annahmen könnte eine Dämpfung dieser Oszillationen durch eine verschieden von unseren Standard-Parametern $\alpha = 1$ und $\beta = 3$ adaptive Wahl von α und β in (5.12) etabliert werden.

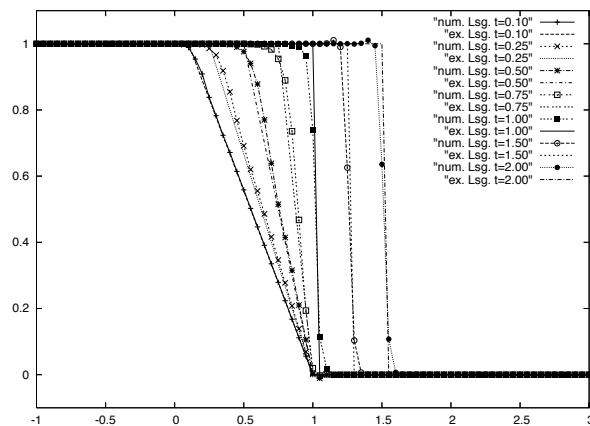


Abbildung 5.15: Numerische Approximation des Cauchy-Problems für (5.30) mit einem 5-Punkte Verfahren vom Lax-Wendroff-Typ (5.36) unter Verwendung von (5.12) mit adaptiver Wahl der Finite Differenz-Parameter α und β .

In Abbildung 5.15 sind die Ergebnisse einer adaptiven Wahl für die Parameter α und β aufgezeigt. Diese basiert auf der Erkennung des Typs der Unstetigkeit in der numerischen Lösung

durch die verwendete Methode selbst und einer für diese Unstetigkeit entsprechenden Ausrichtung von (5.12) durch Variation von α und β . Wir erkennen, dass diese Methode rein von zweiter Ordnung eine deutliche Verringerung der Oszillationen in der numerischen Lösung mit sich bringt. Darüber hinaus ist eine bessere Approximation der numerischen Lösung an die exakte Lösung als im Fall der Methode mit räumlicher Ableitungs-Approximation von erster Ordnung, wie in Abbildung 5.12 gezeigt, festzustellen.

Literaturverzeichnis

- [1] ADAMS, R. A., *Sobolev Spaces*, Academic Press, New York, San Francisco, London, 1975.
- [2] AGRATINI, O., *On approximation of functions by positive linear operators*, Proc. of ICMI 45, Bacau, (2006), pp. 17–28.
- [3] ALTOMARE, F. UND CAMPITI, M., *Korovkin-type Approximation Theory and its Applications*, Studies in Mathematics vol. 17, Walter de Gruyter, Berlin, New York, 1994.
- [4] BAUER, H., *Wahrscheinlichkeitstheorie und Grundzüge der Maßtheorie*, Walter de Gruyter, Berlin, 1968.
- [5] BOHMAN, H., *On approximation of continuous and of analytic functions*, Arkiv för Matematik, 2 (1951), pp. 43–56.
- [6] DEVORE, R. A., *The Approximation of Continuous Functions by Positive Linear Operators*, vol. 293 of Lecture Notes in Mathematics, Springer, Berlin, Heidelberg, New York, 1972.
- [7] FEISTAUER, M., FELCMAN, J. UND STRAŠKRABA, I., *Mathematical and Computational Methods for Compressible Flow*, Oxford University Press, Oxford, 2003.
- [8] GAUSS, C. F., *Theoria motus corporum coelestium in sectionibus conicis solem ambientium*, 1809.
- [9] GODLEWSKI, E. UND RAVIART, P.-A., *Numerical Approximation of Hyperbolic Systems of Conservation Laws*, Springer Verlag, New York, Berlin, Heidelberg, 1996.
- [10] GOLUB, G. H. UND VAN LOAN, C. F., *Matrix Computations - Second Edition*, The John Hopkins University Press, Baltimore, London, 1989.
- [11] HÄMMERLIN, G. UND HOFFMANN, K.-H., *Numerische Mathematik*, Springer, 1994.
- [12] KNOOP, H.-B. UND POTTINGER, P., *Ein Satz vom Korovkin-Typ für C^k -Räume*, Mathematische Zeitschrift, 148 (1976), pp. 23–32.
- [13] ———, *On simultaneous approximation by certain linear positive operators*, Arch. Math., 48 (1987), pp. 511–520.
- [14] KOROVKIN, P. P., *On convergence of linear positive operators in the space of continuous functions*, Dokl. Akad. Nauk SSSR, 90 (1953), pp. 961–964.
- [15] ———, *Linear Operators and Approximation Theory*, International Monographs on advanced Mathematics & Physics, Hindustan Publishing Corp. (India), Dehli, 1960.

- [16] KUNLE, M., *Entwicklung und Untersuchung von Moving Least Square Verfahren zur numerischen Simulation hydrodynamischer Gleichungen*, Dissertation, Fakultät für Physik, Eberhard-Karls-Universität zu Tübingen, 2001.
- [17] LANCASTER, P. UND ŠALKAUSKAS, K., *Surfaces generated by moving least squares methods*, Mathematics of Computation, 37 (1981), pp. 141–158.
- [18] LEVEQUE, R. J., *Numerical Methods for Conservation Laws. ETH Zürich Lectures in Mathematics*, Birkhäuser, Basel, 1992.
- [19] LORENTZ, G. G., *Approximation of Functions*, Chelsea Publishing Company, New York, 1986.
- [20] MUÑOZ-DELGADO, F. J., RAMÍREZ-GONZÁLEZ, V. UND CÁRDENAS-MORALES, D., *Qual. Korovkin-Type Results on Conservative Approximation*, J. Approx. Theory, 94 (1998), pp. 144–159.
- [21] POPOVICIU, T., *Asupra demonstratiei teoremei lui Weierstrass cu ajutorul polinoamelor de interpolare*, Editura Academiei Republicii Populare Romane, (1951), pp. 1664–1667.
- [22] ———, *On the proof of Weierstrass' theorem using interpolation polynomials (translated by Daniela Kasó)*, East Journal on Approximations, 4 (1998), pp. 107–110.
- [23] RICHTMYER, R. D. UND MORTON, K. W., *Difference Methods for Initial-Value Problems – Second Edition*, Interscience Publishers, New York, London, Sidney, 1967.
- [24] RUNGE, C., *Über empirische Funktionen und die Interpolation zwischen äquidistanten Ordinaten*, Zeitschrift für Mathematik und Physik, 48 (1901), pp. 224–243.
- [25] SENDOV, B. UND POPOV, V., *On the convergence of derivatives of linear positive operators (in russian)*, Acad. Bulgare Sci., 22 (1969), pp. 507–598.
- [26] SHEPARD, D., *A two-dimensional interpolation function for irregularly-spaced data*, Proc. ACM National Conference, (1968), pp. 517–524.
- [27] SONAR, TH., *Difference operators from interpolating moving least squares and their deviation from optimality*, ESAIM: M2AN, 39 (2005), pp. 883–908.
- [28] SONAR, TH., NETUZHYLOV, H., UND YOMSATIEANKUL, W., *Finite difference operators from moving least squares interpolation*, ESAIM: M2AN, 41 (2007), pp. 959–974.
- [29] YOSIDA, K., *Functional Analysis*, Springer, Berlin, 1995.

Lebenslauf

Allgemeine persönliche Informationen

Name	Oliver Nowak
Geburtsdatum	30. Juli 1980
Geburtsort	Helmstedt
Staatsangehörigkeit	deutsch
Familienstand	ledig

Schulische Ausbildung

1986 – 1990	Grundschule Lessingstraße, Helmstedt
1990 – 1992	Orientierungsstufe Amtsgasse, Helmstedt
1992 – 1996	Beireis-Realschule, Helmstedt
1996 – 1999	Gymnasium am Bötschenberg, Helmstedt
03.07.1999	Abitur

Wehrdienst

09 1999 – 06 2000	Grundwehrdienst beim Instandsetzungsbataillon 3 in Celle
-----------------------	---

Universitäre Ausbildung

10 2000 – 09 2005	Studium der Mathematik mit Nebenfach Informatik an der TU Braunschweig
27.09.2002	Vordiplom in Mathematik
28.09.2005	Diplom in Mathematik

Wissenschaftliche Tätigkeiten

Seit 10 2005	Wissenschaftlicher Mitarbeiter am Institut Computational Mathematics der TU Braunschweig
----------------	--

